Productivity Effects of Internationalisation through the Domestic Supply Chain*

Bruno Merlevede[†] Angelos Theodorakopoulos[‡]

May 2021

Abstract

This paper investigates productivity effects for a given firm resulting from the import or export of intermediate inputs by domestic upstream and downstream industries. Using manufacturing firms in 19 EU countries over the period 2000-2014, we find that domestic access to intermediate inputs that are also exported leads to higher levels of revenue productivity. The effect appears as more prominent for firms with non-foreign ownership and in relatively downstream, low-tech, or labour intensive industries. Subsequent exploration of mechanisms uncovers patterns consistent with learning by exporting on the part of upstream supplying industries that generates positive productivity spillovers to downstream firms.

Keywords: Importing, Exporting, Intermediate Inputs, Supply Chain, Total Factor Productivity, Learning

JEL classification: D22, D24, D57, D83, F14, L25

*We thank participants at the U4-Workshop Oct. 2014, SMYE May 2015, Ecores Summer School May 2015, ETSG Sept. 2015, ITSG Dec. 2015, Ghent University Internal Seminar Oct. 2015, 8th FIW Dec. 2015, FREIT Jun. 2016, CIE Jun. 2016, and GTD Jun. 2016 for their comments and suggestions. Special thanks are extended to Dan Ackerberg, Andrea Ariu, Sotiris Blanas, Matilde Bombardini, Douglas L. Campbell, Davide Castellani, Ruben Dewitte, Gerdie Everaert, Rebecca Freeman, Dirk Van de gaer, Jan Kiviet, Joep Konings, Sebastian Kripfganz, Kun Li, Bart Los, Inmaculada Martínez-Zarzoso, Klaas Mulier, Gianmarco Ottaviano, Nina Pavcnik, Glenn Rayp, David A. Rivers, Mark J. Roberts, Ilse Ruyssen, Johannes Schmieder and Samuel Standaert. We also thank the editor, Marco Del Negro, and three anonymous referees for their guidance and comments. This paper previously circulated under the names "Productivity Effects of Internationalisation through the Domestic Supply Chain: Evidence from Europe" and "Productivity Effects from Inter-industry Offshoring and Inshoring: Evidence from Belgium."

[†]Department of Economics, Ghent University, Ghent, Belgium; e-mail: `bruno.merlevede@ugent.be`

[‡]Corresponding author: University of Oxford, Oxford Martin School, 34 Broad Street, Oxford OX1 3BD, United Kingdom; e-mail: `angelos.theodorakopoulos@oxfordmartin.ox.ac.uk`

1 Introduction

How do the trade-related activities of a firm's domestic suppliers and/or customers impact its productivity? While many studies have shown that exporting and importing enhance firms' performance, less is known about productivity spillovers to their domestic supply chain partners.

On the one hand, a large literature has examined the effects of export behaviour on firm productivity. Termed "learning by exporting", productivity enhancements of this sort can be explained by several mechanisms. Some of these include: product innovation; process upgrading; and improvements in technical standards, managerial practices, and inventory techniques.[1] On the other hand, "learning by importing" enhances firm productivity through access to foreign-sourced intermediates, which are differentiated from those available on the domestic market. In turn, these inputs can lead to knowledge transfer, access to more varieties, and cost savings from process and product innovation. Learning by importing can also boost the efficiency of the domestic component of production processes by complementing domestic inputs.[2]

In the domestic economy, firms (which might also engage in international trade) are linked to these exporters and importers through the domestic supply chain. They may source inputs domestically from upstream suppliers. They may also provide inputs domestically to downstream customers. When either (or both) is the case, do the productivity benefits of internationalisation experienced by firm's local suppliers and/or customers ripple through the domestic supply chain? This question remains largely unanswered[3] (Shu and Steinwender 2019) and is the topic of our analysis.

Figure 1 illustrates the four potential channels we consider whereby internationalisation could affect firms' performance through the domestic supply chain. Upstream importing and upstream exporting start from a given local firm i in industry j that sources inputs domestically from upstream suppliers u. Upstream importing is the channel through which pass-on productivity effects from domestic suppliers that import intermediates could occur. Upstream exporting considers the pass-on productivity effects from domestic suppliers' exporting activity. On the flip side, downstream importing and downstream exporting consider the same given local firm i in industry j that may also supply inputs domestically to downstream customers d. Downstream importing refers to the pass-on productivity effects from the importing activity of downstream customers to their local suppliers. Finally, downstream exporting captures pass-on productivity effects which are associated with supplying domestic clients that export intermediates. As such, we shed light on potential productivity effects to a given firm associated

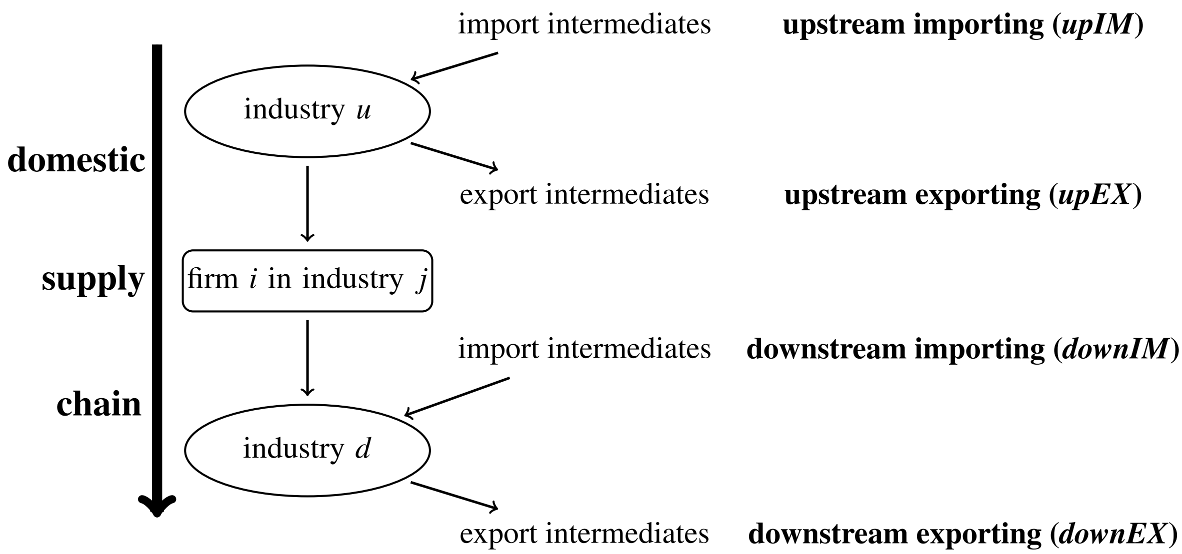
[1] Clerides et al. (1998); Van Biesebroeck (2005) and De Loecker (2007) examine learning by exporting effects. Prominent papers which can explain these effects include those by Bustos (2011) and Lileeva and Trefler (2010).

[2] Amiti and Konings (2007) and Kasahara and Rodrigue (2008) examine learning by importing effects. Papers such as those by Markusen (1989), Grossman and Helpman (1991), Goldberg et al. (2010), Bas and Strauss-Kahn (2015), Halpern et al. (2015) and Antràs et al. (2017) present explanations of these effects.

[3] A notable exception is Blalock and Veloso (2007), who use Indonesian firm data to examine how the importing activity of downstream firms impacts the productivity of their domestic suppliers. They find a productivity boost among domestic suppliers due to increased competition from abroad.

with the internationalisation behaviour of its domestic suppliers and customers.[4]

Figure 1: Inter-industry importing and exporting for firm i in domestic industry j



As in most empirical research which addresses firm performance, productivity measurement is a core component in our analysis. The approach we use follows that of Gandhi, Navarro, and Rivers (2020) (herein GNR). Their estimation method of gross output production functions with at least one flexible input controls for value-added bias found in other estimators and endogeneity, as well as allows for the inclusion of firm fixed effects.

Beyond controlling for the presence of endogeneity when estimating the production function—i.e. firms know their productivity when choosing inputs—it is also important to address potential endogeneity of inter-industry importing and exporting. Endogeneity of this sort could arise due to omitted variable bias. For example, unobserved characteristics such as improvements in regional infrastructure could both have a positive effect on a given firm's productivity and facilitate the exchange of inputs with domestic suppliers or customers that trade internationally. We address endogeneity of this sort in our baseline specification through simultaneously introducing: time-varying controls and fixed effects at different levels of aggregation; firm fixed-effects; and an internal instrumental variable (IV) strategy based on deeper lags of the relevant variables (*à la* Blundell and Bond 1998).

To further demonstrate the causal interpretation of our results, we complement our analysis by applying an external IV strategy. The goal is to verify that our main results are not driven by unobserved domestic shocks which are persistent over time and thus correlated with our variables of interest. For example, a targeted industrial policy that improves efficiency in a given industry and leads to increased demand for local inputs could induce local suppliers to import more or export less themselves. To guard against this type of endogeneity, we exploit

[4]Given the data at hand, our notion of importing and exporting refers to the exchange of intermediate inputs only and abstains from final goods trade.

two identification ideas from the literature. The first follows Autor et al. (2013) and Dauth et al. (2014) who draw on variation from trade flows of other high-income countries as instruments for domestic import and export exposure. The second exploits the arrival of China in the world market as a quasi-natural event. In both cases, identification comes from shocks abroad that drive the internationalisation behaviour of industries linked through the domestic supply chain, while excluding domestic shocks in the industry of the firm considered.

Throughout this analysis, we combine a rich micro-level dataset for firms in the manufacturing sector with annual country-industry data from input-output (IO) tables for 19 EU countries. Our firm-level dataset contains all balance sheet information necessary for the estimation of the firms' production function (and hence productivity). The IO-tables are then used to construct measures of inter-industry importing and exporting intensities which vary at the industry-country-year level. Although the level of aggregation of the latter measures is a clear limitation of our analysis, our consideration of a broad set of countries lends itself to strong external validity of our results, as the 19 EU countries considered differ across various dimensions: geography; economic development; types of institutions; trade openness/integration; etc.

Our results show that sourcing from domestic industries that also export intermediates leads to higher productivity levels of a given firm (termed as upstream exporting). This prevails as the only robust channel through which inter-industry productivity effects occur.[5] Moreover, we find that a one standard deviation increase in upstream exporting is associated with a productivity increase of 0.75% in the short run and 2.58% in the long run. This result could potentially explain the finding of Amiti and Konings (2007), where, on average, even non-importing firms benefit from tariff reductions.

Our results indicate that the domestic supply chain can facilitate indirect productivity effects from internationalisation. Exploring further heterogeneity, we find that effects tend to be stronger for firms that are less likely to have previous international involvement. Specifically, our results suggest that the effect is more prominent for firms without foreign ownership links and for firms in relatively downstream, low-tech, or labour intensive industries (or combinations thereof).

The mechanisms we have in mind to explain this finding bear close resemblance to vertical productivity effects from foreign direct investment (FDI), as in Javorcik (2004) and Blalock and Gertler (2008). In our case, however, the conduit is exporting by vertically related firms. Specifically, it has been shown that firms can learn from performing tasks such as organisational restructuring, network sharing, managerial practices, and transferring knowledge (Arrow 1962; Stokey 1988; Parente 1994; Jovanovic and Nyarko 1996). Learning can also be relationship-specific when production requires the coordination of inputs from multiple firms (Kellogg 2011). In this paper, we associate the learning process with sourcing intermediate inputs from upstream sectors that also export.

[5] Our results, if anything, weakly support the channel explored in Blalock and Veloso (2007), i.e. firms that supply customers in downstream industries which also export intermediates see productivity enhancements. However, they focus on one specific developing country while we use a multi-country dataset.

Lastly, to explore the significance of learning as an underlying mechanism, we exploit additional sources of variation in the data stemming from the development levels of the countries in which firms and their trading partners are located. We find that only firms in developing/transition economies, i.e. in Central Eastern Europe, tend to benefit from productivity improvements from upstream exporting. Most interestingly, these effects exist only when the trading partner of the upstream supplier is a technologically developed economy. Intuitively, firms that are further away from the technological frontier are actually the ones to benefit the most from increased access to foreign markets through the domestic supply chain. Consistent with the learning by exporting literature, our results suggest that learning mechanisms play a crucial role in explaining how the productivity effects of internationalisation ripple through the domestic supply chain.

The remainder of this paper is organised as follows. In Section 2 we discuss inter-industry importing and exporting and present how the proxies used throughout our analysis are constructed. In Section 3 we describe the empirical model and how it's estimated. Section 4 describes the data and Section 5 presents results. It further explores various forms of heterogeneity and potential mechanisms at play. Finally, Section 6 concludes.

2 Inter-industry Importing and Exporting

To measure inter-industry importing and exporting activities, as depicted in Figure 1, we construct proxies at the industry-country-year level using Input-Output (IO) tables from the World Input-Output Database (WIOD).

A firm is confronted with downstream importing when domestic clients start to import intemediates previously sourced at home (see Figure 1). Downstream importing thus captures downstream demand side shocks or import competition. Firms then face foreign competition to supply differentiated intermediate inputs to domestic downstream industries. To survive and remain competitive, these firms need to reduce costs and improve productivity. Should the intermediates be complements to those imported by domestic downstream firms, productivity effects may result from upgrading production processes. This is done to match specificities and quality standards of the complementary intermediate inputs used in the production process of domestic downstream firms (see Blalock and Veloso 2007). To capture downstream importing, we build on a measure introduced by Merlevede and Michel (2020):

$$downIM_{jct} = \sum_{d \neq j} \theta_{jdct} \Phi_{jdct} \tag{1}$$

where Φ_{jdct} is the import intensity in industry d of products that industry j supplies to domestic downstream industry d in country c in year t; it is calculated as the share of imported intermediates in total intermediates sourced from j by d. Φ_{jdct} thus measures the importing behaviour of domestic downstream industry d that directly affects (local firms in) industry j. We obtain a value of Φ for all possible industry pairs involving a given industry j and use θ_{jdct} to generate a single value for $downIM_{jct}$ as a weighted average of Φs. More precisely, θ_{jdct} is the proportion of industry j's domestic supply sold to downstream industry d within country c in year t. θs are calculated from the domestic IO-table.

Similarly, we define 'upstream exporting' for a given firm as sourcing intermediates from (a firm in) a domestic industry that also exports the same intermediates (see Figure 1). For instance, think of a domestic firm in the Manufacture of Computer, Electronic and Optical Production industry that supplies intermediate inputs to a local firm in the Manufacture of Electrical Equipment industry. When the former also exports these products for intermediate use abroad this may affect the productivity of the domestic client because of the availability of higher quality domestic inputs. Potential productivity effects may also extend to more indirect mechanisms such as diffusion of management practices, benefits from international networking, and organisational restructuring. To analyse such effects we define upstream exporting as:

$$upEX_{jct} = \sum_{u \neq j} \zeta_{juct} \Psi_{juct} \tag{2}$$

where Ψ_{juct} measures the exporting intensity of domestic upstream industry u with respect to products that industry j uses as intermediates. Export intensities are calculated as the share of exported intermediates in the total amount of intermediates delivered by u. We average Ψs over domestic partner industries u using ζ_{juct}, defined as the proportion of industry j's domestically sourced intermediates from upstream industries u within country c at time t, to obtain $upEX_{jct}$.

In addition to the aforementioned direct channels which are based on intermediates common to industry j and its domestic partner industries, we analyse two further forms of indirect internationalisation through domestic supply chain participation. This includes potential productivity effects from upstream importing and downstream exporting, where no product-specific links are at play. However, importing in a previous stage or exporting in a subsequent stage of the domestic supply chain may also result in indirect productivity spillover effects similar to those above. Upstream importing is a supply side effect originating from the import of intermediate inputs by a firm's domestic suppliers (see Figure 1). In this case, the learning mechanisms rely on knowledge diffusion from upstream to downstream domestic industries (Grossman and Helpman 1995; Coe and Helpman 1995; Connolly 2003). Downstream exporting effects may result from the demand for increased quality of intermediate inputs from export oriented domestic downstream clients. To analyse such effects we define upstream importing as:

$$upIM_{jct} = \sum_{u \neq j} \zeta_{juct} \Omega_{uct} \tag{3}$$

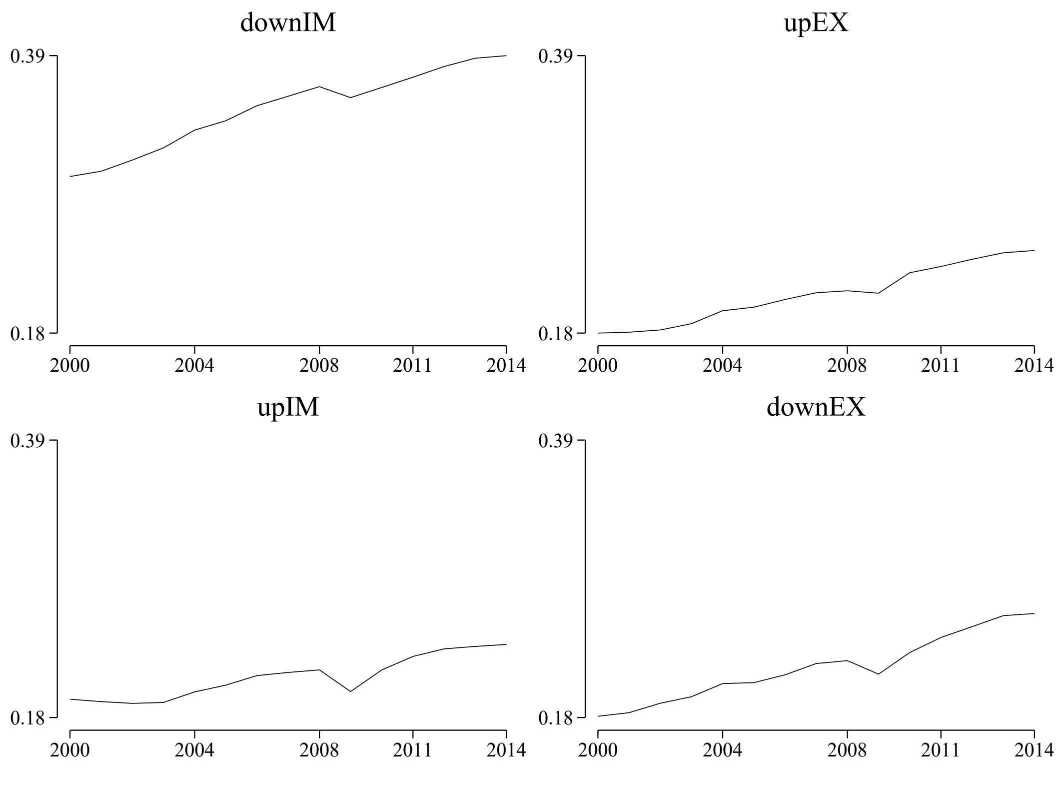
where ζ_{juct} is again the proportion of industry j's intermediate inputs sourced from domestic upstream industries u at time t. Ω_{uct} is the 'overall' import intensity in industry u, averaged over all products since in this case there is no direct product link between industries j and u. Import intensities are calculated as the share of imported intermediates in total intermediates used by u.

Downstream exporting is then defined as:

$$downEX_{jct} = \sum_{d \neq j} \theta_{jdct} \Theta_{dct} \tag{4}$$

where Θ_{dct} measures the export intensities of domestic downstream industries d and θ_{jdct} is defined as above. Export intensities are calculated as the share of exported intermediates in the total amount of intermediates sold by d. For robustness, we also consider measures in (1)-(4) where we fix weights in the initial period. Additionally, we consider alternative measures of (1)-(4) which incorporate the diagonal elements when $d = j$ and $u = j$ in the summation, and which represent the intra-industry importing and exporting intensities. Next, we include alternative broad measures of (1) and (2) using the 'overall' import (Φ_{dct}) and export (Ψ_{uct}) intensities, instead of those based solely on the product links between downstream (Φ_{jdct}) and upstream (Ψ_{juct}) industries, respectively.

Figure 2: Inter-industry importing and exporting by year (averaged over countries and industries)

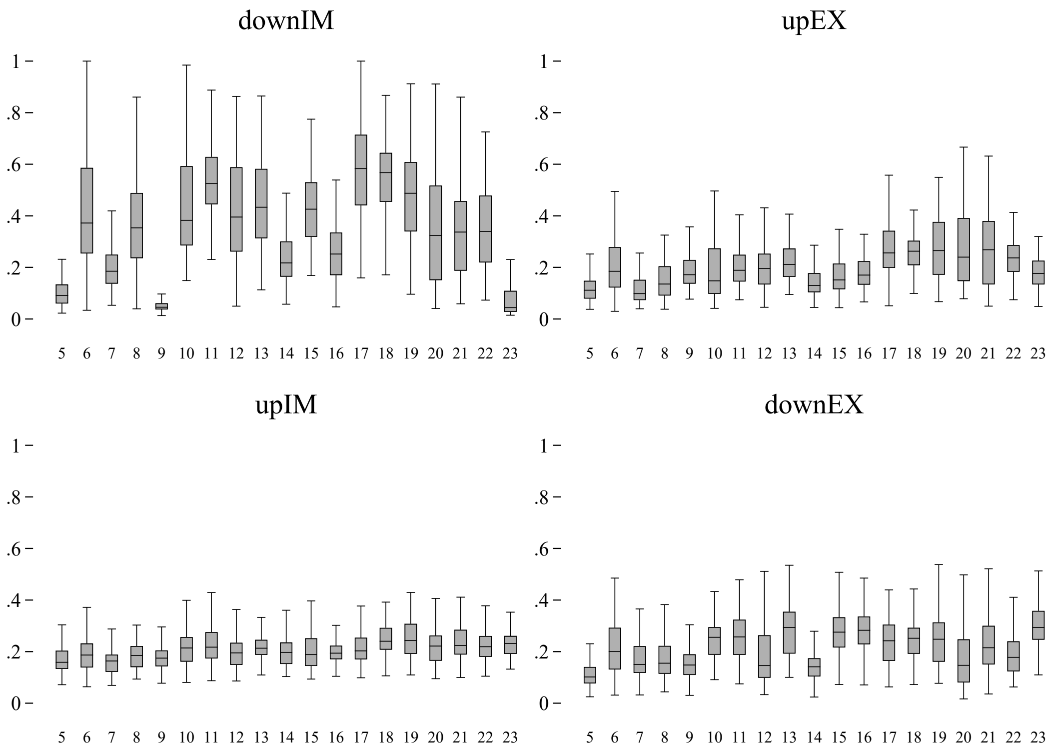


Source: Authors' calculations based on WIOT.

Furthermore, our empirical model discussed in section 3 includes intra-industry import (IM) and export (EX) intensities of intermediate inputs among the explanatory variables. These are important control variables for two principal reasons. First, our firm-level data do not allow us to determine firm-specific import and export intensities and the industry-level intensities are an aggregate over firms with and without internationalisation activities. Second, our measures in (1)-(4) exclude intra-industry supply of intermediates in order to capture pure inter-industry effects. Since the industry classification in the IO-tables is fairly aggregated, within-industry supply chain relations are likely to exist and will also be reflected in industry-level import and export intensities. This makes the intra-industry variables important controls, but their interpretation is hampered by the fact that they reflect a net outcome of different mechanisms.

For all four measures, industries with larger values are those facing relatively more downstream/upstream importing/exporting. Figure 2 presents the average trend for each of the inter-industry variables in equations (1)-(4). For the case of the EU we find a clear upward trend across all measures with a justified decrease in the year following the 2008 financial crisis. The clear overall upward trend fits with the EU's economic history characterised by increasingly integrated economies which are heavily oriented towards intra-bloc trade in intermediates.

Figure 3: Inter-industry importing and exporting by industry



Source: Authors' calculations based on WIOT.
Notes: Let x represent the variable of interest. The upper, middle and lower hinge of the box represents the 25th ($x_{[25]}$), 50th ($x_{[50]}$) and 75th ($x_{[75]}$) percentile, respectively. Define $x_{(i)}$ as the ith ordered value of x. The upper adjacent line has a value $x_{(i)}$ such that $x_{(i)} \leq U$ and $x_{(i+1)} > U$, where $U = x_{[75]} + 1.5(x_{[75]} - x_{[25]})$. The lower adjacent line has a value $x_{(i)}$ such that $x_{(i)} \geq L$ and $x_{(i+1)} < L$, where $L = x_{[25]} - 1.5(x_{[75]} - x_{[25]})$.

We observe heterogeneity in the measures originating from the country and/or industry dimension. Figure 3, for example, contains boxplots of the values for each of the variables by industry (see also Figure D.1 in Online Appendix D). We observe substantial variation across industries: industry 17, Manufacture of computer, electronic and optical products, together with industry 18, Manufacture of electrical equipment, score high on all four variables. Industry 5, Manufacture of food products, beverages and tobacco products, closely followed by industry 7, Manufacture of wood and products of wood and cork, except furniture; etc., are generally confronted with the lowest values for inter-industry internationalisation. Similarly, at the country dimension we find Hungary and Estonia to score the highest while Spain and Italy are facing the lowest values across all four measures (see Figure D.2 in Online Appendix D). Table 1 in the data section contains further summary statistics.

3 Empirical Methodology

In this section we first provide a brief overview of the different production function estimation methods that control for the endogeneity of production inputs. Next, we present a baseline empirical model for our variables of interest. Continuing, we describe relevant details and steps followed when bringing our baseline specification to the data under an augmented version of the production function estimation procedure. Finally, we discuss how we account for potential endogeneity of the variables of interest and describe how we implement cluster-robust inference.

3.1 Total Factor Productivity

To analyse potential productivity effects of inter-industry importing and exporting, we start by specifying a production function $Y_{it} = F_j(K_{it}, L_{it}, M_{it}; \alpha_j)\mathrm{e}^{\omega_{it}+\varepsilon_{it}}$ for firm $i = 1, \ldots, N$ (in industry j and country c) at time $t = 1, \ldots, T$. We consider an industry j-specific production technology F_j governed by the set of parameters α_j and a Hicks-neutral total factor productivity ω_{it} (TFP). In logs, the production function to be estimated takes the following form:

$$y_{it} = f_j(k_{it}, l_{it}, m_{it}; \alpha_j) + \omega_{it} + \varepsilon_{it} \tag{5}$$

where y_{it}, k_{it} and m_{it} are log values of deflated sales, tangible fixed assets and material costs, respectively, and l_{it} is the log of the total number of employees. ω_{it} is unobserved to the econometrician but known to the firm. Identically and independently distributed (i.i.d.) shocks which occur after the firm's decision about which inputs to use are picked up by ε_{it}.

Traditionally, the applied production function estimation literature has employed structural approaches to estimate (5). Two primary approaches include dynamic panel methods (Arellano and Bond 1991; Blundell and Bond 1998, 2000) and proxy variable methods (Olley and Pakes 1996; Levinsohn and Petrin 2003; Ackerberg et al. 2015). In these approaches, the main focus has been to solve for endogeneity, also known as 'simultaneity' or 'transmission bias.' Such bias originates from the fact that firms know their productivity level when they decide on which inputs to use (Marschak and Andrews 1944; Griliches and Mairesse 1999). Proxy variable methods have dominated in the literature given dynamic panel methods' weak performance both at a theoretical and empirical level. Despite their popularity and prevalence, proxy variable methods suffer from identification issues when the production function contains at least one flexible input such as materials. These issues have been highlighted by Mendershausen (1938); Marschak and Andrews (1944); Bond and Söderbom (2005); Ackerberg et al. (2015) and formalised by GNR.

To overcome this issue and identify the effects of interest, we instead follow the estimation approach put forth by GNR. They propose a simple, non-parametric estimator for production functions with at least one flexible input. In this method, identification is established by exploiting information in the first order condition with respect to the flexible input from the firm's static profit maximisation problem. In line with most of the proxy variable methods, the GNR procedure follows two steps and allows us to both estimate the production function and identify the productivity effects from inter-industry importing and exporting. A second benefit is that, as opposed to standard proxy variables methods, the GNR framework allows for the inclusion of firm fixed effects which become particularly important for our analysis (see Section 3.4). Online Appendix A outlines the assumptions and steps followed under this identification strategy.

Note that TFP is not identical to disembodied technological change, known as the 'Solow Residual' (Solow 1957). Here TFP also includes the impact of inputs that are not explicitly measured (e.g. human capital). Further, our TFP estimates are revenue based as we do not observe physical output, but only monetary values deflated at the industry-country-year level. Results should be interpreted bearing this caveat in mind (Klette and Griliches 1996).

3.2 Effects of Inter-industry Importing and Exporting on TFP

We now specify how we model the effects of inter-industry importing and exporting on TFP by allowing the measures of interest to shift the productivity path, ω_{it}. Specifically, we introduce the relevant proxies and other control variables in the law of motion of productivity—as in Aw et al. (2008); Kasahara and Rodrigue (2008); Doraszelski and Jaumandreu (2013) and De Loecker (2013)—which we substitute in (5), resulting in the following dynamic representation:[6]

$$y_{it} = f_j(k_{it}, l_{it}, m_{it}; \alpha_j) + \rho_\omega \omega_{it-1} + \rho_p proxies_{jct-1} + \rho_x X_{it-1} + \phi_{fe,t} + \phi_i + \xi_{it} + \varepsilon_{it} \qquad (6)$$

where $proxies_{jct-1}$ represents the vector of inter- and intra-industry importing and exporting proxies at the industry-country-year level (jct); X_{it-1} is a vector of additional firm-level controls; $\phi_{fe,t} = \left(\phi_t, \phi_j t, \phi_c t\right)$ is a vector of time-varying fixed effects, including a set of dummies for time fixed effects ϕ_t, and a set of industry and country specific linear time trends $\phi_j t$ and $\phi_c t$,

[6] As formally derived by De Loecker (2013), this one-stage formulation accounts for the inconsistencies and biases that arise from a two-stage specification where: (i) in a first stage, a TFP estimate is obtained when the variables of interest are not taken into account in the law of motion of TFP; and (ii) by a second stage where the estimated TFP from the first stage as a dependent variable is regressed on the variables of interest.

respectively; ϕ_i is a set of dummies for firm fixed effects; and ξ_{it} is an i.i.d. error term capturing innovations to productivity.[7] Therefore, (6) is our baseline specification (henceforth Baseline) that we bring to the data following the estimation procedure discussed in the next section.

3.3 Estimation

To maintain sufficient variation for the consistent estimation of both the production function $\big(f_j(\cdot)\big)$ and the effects of interest (ρ_p) in (6), we pool firms across all manufacturing industries and European countries in the data. This procedure is in line with estimators of dynamic panel models with a short time dimension and individual specific effects (Blundell and Bond 2000). In principle, one could run the estimation with more granular groups, e.g. by pooling firms at the country or country-industry level separately. However, this would result in sample selection issues since certain groups would not have a sufficient number of firm-year observations to identify the industry-specific production function. This is especially true for smaller countries.[8] In addition, given the aggregate nature of our variables of interest (i.e. country-industry-year specific), a more granular estimation would narrow down their effective identifying variation, jeopardising the estimation precision.[9] To guard against such limitations, our estimation relies on observations pooled across all manufacturing industries and allows for industry-specific production technologies to be estimated jointly. Acknowledging that this approach might mask potential heterogeneous effects across countries and/or industries, we later try to uncover these with additional checks and meaningful sample splits.

For the production technology we rely on a Cobb-Douglas functional form: $f_j(\cdot) = \sum_j \sum_v \alpha_v^j v_{it}$, where $v = \{k, l, m\}$ and the sum over j allows for industry-specific production technologies to be jointly estimated when pooling the data across all industries as discussed above.[10] This specification is the simplest and most commonly used in the literature, albeit at the expense of restricting the elasticities of substitution between inputs to unity. However, given the parameter space and data constraints in mind, its simplicity allows us to control for additional dimensions, i.e. growth differentials in production technologies across industries (j) and countries (c), without depleting the degrees of freedom and impeding the estimation routine.[11]

Following the dynamic panel literature, GNR augment the second step of their estimation

[7] The mean productivity (ρ_0) and the constant of the production function $(f(\cdot))$—which are not separately identified—are subsumed into the firm fixed effects(ϕ_i).

[8] To name a few, these include: manufacture of coke and refined petroleum products for most countries in the sample; manufacture of basic pharmaceutical products and preparations for Croatia, Slovenia and Slovakia; and manufacture of other transport equipment for Austria, Estonia and Hungary; etc.

[9] For example, by using country-industry level information we can only rely on inter-temporal changes, which would be hard to disentangle from aggregate trends and business cycle co-movements.

[10] Since the production function is industry specific (j), it is imperative to use industry fixed effects (ϕ_j). However, for notational simplicity, in (6), all time invariant controls (including ϕ_j) are subsumed in ϕ_i.

[11] Alternatively, one could allow for more flexible substitution patterns between inputs, i.e. translog functional form, and/or production technologies that vary both across industries and countries. However, such choices come with typical trade-offs faced by empirical researchers: increased parameter space; insufficient number of observations for certain industry-country pairs; and computationally intensive estimation routines.

procedure to control for firm fixed effects by: (a) first-differencing them out; and (b) instrumenting with lagged levels of the now endogenous (first-differenced) variables *à la* Arellano and Bond (1991) (see Online Appendix A.2.1). However, as shown by Blundell and Bond (1998), first differencing in a dynamic panel setup performs poorly when ω_{it} is close to a random walk because of weak instruments causing large finite sample bias. Therefore, to reduce such biases, we further augment the GNR estimation procedure, borrowing from the "System GMM" estimator developed by Blundell and Bond (1998) and outlined by Arellano and Bover (1995). More concretely, we simultaneously estimate the first-differenced and levels equations by distinctly instrumenting with the lagged levels and lagged first-differences of the endogenous variables, respectively (see Online Appendix A.2.2). This will become our procedure for estimating baseline specification (6) upon which economic inference will be made.

3.4 Endogeneity

Until now, we have discussed how to control for the endogeneity of inputs in a production function setup. However, there is also the concern about potential endogeneity of our variables of interest. This can arise for two main reasons: simultaneity and/or omitted variable bias. To alleviate such concerns, we first discuss the relevant steps taken to address this endogeneity in our baseline estimation approach and through robustness checks. In turn, we complement the baseline approach by proposing an alternative IV strategy with external instruments.

3.4.1 Baseline

Endogeneity can arise if our variables of interest and TFP are simultaneously determined (simultaneity bias). Our baseline specification addresses this through the assumption that the activity of upstream or downstream firms is not immediately observed; it is gradually explored by the firm. This is a reasonable assumption given that it takes time for firm decisions to ripple through the domestic supply chain. Importantly, this assumption addresses simultaneity bias because it implies a delay in the transmission of productivity effects. As such, we use one year lagged proxies to model such sluggishness. In turn, current TFP is unlikely to affect lagged values of inter-industry importing and exporting.[12]

Nevertheless, firm performance has been empirically shown to be persistent over time (Syverson 2011). To the extent that TFP adjusts dynamically, endogeneity could arise if the lagged proxies of interest are also affected by lagged TFP. We thus include lagged TFP as a regressor in our baseline model. This eliminates any endogeneity concerns by controlling

[12]Consistent with the timing assumption for material inputs used in estimating TFP, intra-industry importing and exporting are expected to contemporaneously affect firm productivity. This is because material inputs are assumed to be flexible. In this case, upon data availability, the decision to import or export is endogenous and should be modelled accordingly. See Görg et al. (2008); Michel and Rycx (2014); and Halpern et al. (2015). Due to a lack of data availability, we cannot directly split firm-level material inputs into a domestic and foreign component—also known as the "static effect" (Kasahara and Rodrigue 2008). However, this should not impose any threat to our research question since we are interested in the future TFP effects from inter-industry importing and exporting.

for potential causal effects running from TFP to our lagged variables of interest through the persistence term.

With regards to omitted variable bias, this can arise if any unobserved determinants of TFP are correlated with our variables of interest (and thus captured in the error term). We alleviate concerns of omitted variable bias in our baseline estimation in three main ways.

First, a given firm might have some impact on the importing and exporting choice of its client or supplier industries through affiliated firms. If unobserved, this would render our proxies endogenous for sufficiently granular firms. To mitigate this concern, we include additional firm-level controls X_{it-1}. In a first instance, we construct two dummy variables to indicate whether the firm owns any domestic subsidiaries or is owned by any other domestic firm. These are used in conjunction with the firm's multinational status (i.e. whether it has a foreign shareholder or a foreign subsidiary). These controls thus break the correlation between our variables of interest and any type of domestic demand or supply chain relationship between parent and affiliate firms.

In addition, the inter-industry proxies are likely to be correlated with intra-industry import and export intensities for two main reasons. First, given the aggregate nature of the IO tables, additional granular inter-industry supply chain relations are likely to exist, but will only be reflected in the intra-industry measures. Second, shocks originating from abroad (e.g. foreign demand shocks) could simultaneously determine both the intra- and inter-industry decision to export and import intermediates. Since intra-industry importing and exporting are known drivers of firm performance, failing to control for them is likely to result in an omitted variable bias. Therefore, the set of proxy variables $proxies_{jct-1}$ used in baseline equation (6) includes the intensities for intra-industry importing and exporting of intermediates (on top of the four inter-industry proxies). As discussed above and in section 2, these proxies are important controls to help pin down identification. However, economic interpretation of these controls is not meaningful since they could confound the net outcome of different channels.

Second, we control for firm fixed effects to account for any unobserved time-invariant heterogeneity, such as a firm's location, firm-to-firm linkages, or more aggregate geographic/industrial traits, etc. If the endogeneity arises because of correlation between unobserved firm characteristics and our variables of interest, controlling for these fixed effects addresses the problem by focusing the analysis on within-firm variation.

On top of this, we also add a rich set of time-varying fixed effects to account for other unobserved factors which could be driving growth performance at the industry and country level. These include: (a) yearly fixed effects; (b) industry-specific time trends; and (c) country-specific time trends. Introducing yearly fixed effects further address omitted variable bias concerns because it allows for secular changes in the economic environment that affect all firms in the same way (e.g. inflation or aggregate spending). Industry-specific time trends address omitted variable bias by proxying for technical progress in each industry. Finally, country-specific time trends alleviate systematic time patterns such as business cycle trends across countries. Failing to control for such aggregate trends could lead to spurious correlation between firm performance and our variables of interest.

Third, we treat our variables of interest as endogenous in our baseline setup, i.e. current values of our proxies are correlated with the error term. However, we assume that any past values remain orthogonal. Therefore, we implement an IV strategy by instrumenting the variables of interest with deeper lags (internal IV)—$(t-2)$ for the first-differenced equation (*à la* Arellano and Bond 1991) and lagged first-differences $(t-1)$ for the equation in levels (*à la* Blundell and Bond 1998; Arellano and Bover 1995).

Nevertheless, one cannot exclude the possibility of unobserved determinants of productivity which could remain in the background. Thus, in the results section we also provide a series of robustness checks that allow us to further mitigate concerns about the presence of endogeneity which could arise if our modelling assumptions from baseline specification (6) were incorrect. For example, the internal IV strategy implicitly assumes the absence of any persistent unobserved TFP determinants which are correlated with our variables of interest. To ensure that this is not the case, we check the sensitivity of our results against any type of business cycle volatility both at the country and industry level by including country-time and industry-time fixed effects.

3.4.2 Baseline with External IV

Beyond the above, further endogeneity could still arise because of any persistent unobserved productivity determinants which drive the export and import decisions of firms in domestic upstream and downstream industries. For instance, persistent domestic demand shocks could jointly drive firm performance and inter-industry importing/exporting behaviour, leading to spurious estimates.

To mitigate this concern, we complement our baseline analysis by replacing the internal IVs with an alternative set of external IVs. The broad intuition of the external IV strategy is that shocks abroad drive the internationalisation behaviour of industries linked through the domestic supply chain. Originating abroad, these shocks are posited to be uncorrelated with domestic shocks in the industry of the firm considered. As such, this variation provides external instruments which are orthogonal to the error term in specification (6). Importantly, as discussed in section 2, intra-industry importing and exporting variables remain paramount controls in our setup. Therefore, the external IVs also include instruments for intra-industry importing and exporting. Instrumenting for these variables is key, as they might suffer from the same sources of endogeneity as inter-industry proxies. Overall, this identification strategy, as used in the literature already, leads us to construct two sets of IVs (described in detail below).

IV1.—The first IV set ($IV1$) relies on the import and export behaviour of the same (upstream and downstream) industries in a group of other similar, non-sample countries. Following Autor et al. (2013), Dauth et al. (2014), and Merlevede and Michel (2020), foreign shocks which drive other countries' imports and exports are unrelated to domestic developments that influence the trading behaviour of domestic upstream and downstream industries.[13] To ensure the relevance of $IV1$, we rely upon a sample of countries with similar levels of technological development and industrialisation to those in our baseline sample.

$IV1$ is constructed in a similar fashion to our variables of interest. We keep the same weights, but replace the importing and exporting intensities of the linked industries in country c with the relevant intensities, calculated based on a group of non-sample countries $\tilde{c}$ (defined below). The instrument for downstream importing can then be computed as follows:

$$downIM_{jct}^{IV1} = \sum_{d \neq j} \theta_{jdct} \Phi_{jd\tilde{c}t} \tag{7}$$

where we select the non-sample countries ($\tilde{c}$) under the requirement that their shocks are not correlated with the domestic shocks that could affect the 19 European countries in our sample. Similar to Dauth et al. (2014), we use a set of developed countries that excludes neighbouring countries and other EU member states in order to alleviate potential endogeneity concerns from business cycle synchronization.[14] A similar logic is then applied to obtain $upEX_{jct}^{IV1}$, $upIM_{jct}^{IV1}$, and $downEX_{jct}^{IV1}$. In parallel, we use this approach to construct the intra-industry IVs IM_{jct}^{IV1} and EX_{jct}^{IV1}. Overall, these inter- and intra-industry instruments are intended to capture exogenous global export and import opportunities.

IV2.—The second IV set ($IV2$) relies on the arrival of China in the world market to instrument for the variables related to importing and exporting intermediate inputs along the domestic supply chain. Using China's emergence as a production hub is seen as an exogenous shock across two dimensions. It drives exports to other non-sample nations and opens opportunities as an importer from non-sample countries.

In line with the argumentation of Autor et al. (2016), China's arrival to the world market offers a unique source of exogenous variation for the following three reasons. First, China's transition from communist isolation rapidly unlocked the country's potential. Second, the timing and extent of this transition was heavily driven by domestic idiosyncratic political factors. Finally, despite China's comparative advantage in labour intensive production, its specialisation patterns have been influenced by political motives to transform the country into a key player globally. These factors led to unprecedented variation in trade and performance across industries of similar factor content. Therefore, trade flows from non-sample countries are, in turn, posited to be uncorrelated with domestic industrial developments in sample countries. Moreover, $IV2$ is relevant given China's dominance in world production, and hence all countries' reliance on it for intermediate inputs.

[13] Specifically, these papers use external IVs that exploit import penetration from China and Eastern Europe.

[14] Given the country availability in WIOD and the development-level ranking in UNCTAD (2019), we use Australia, Canada, Japan, South Korea, Taiwan, and United States. Despite not being a high-income country, Taiwan ranks among the best countries in terms of technological development and industrialisation (OECD 2019).

Precisely, the instrument for downstream importing is computed as follows:

$$downIM_{jct}^{IV2} = \sum_{d \neq j} \theta_{jdct} Exports_{jdt}^{China \to \tilde{c}} \tag{8}$$

where $\tilde{c}$ now refers to non-sample countries available in WIOD that are not members or bordering the EU.[15] $upIM_{jct}^{IV2}$ follows straightforwardly.

In a similar vein, upstream exporting is written as:

$$upEX_{jct}^{IV2} = \sum_{u \neq j} \zeta_{juct} Imports_{jut}^{China \leftarrow \tilde{c}} \tag{9}$$

where $\tilde{c}$ refers to the same set of non-sample countries defined above. Downstream exporting IV ($downEX_{jct}^{IV2}$) is computed in a similar way. Following the reasoning described for $IV1$, intra-industry importing and exporting are instrumented using Chinese exports to and imports from non-sample countries (IM_{jct}^{IV2} and EX_{jct}^{IV2}).

To increase the strength of the identifying variation, we combine both $IV1$ and $IV2$ in the IV set used to generate the moment conditions that identify the model. Finally, as the US comprises a disproportionately large share of the world economy, it is possible that US shocks and domestic shocks in our 19 European country sample have stronger correlation. To account for this, we calculate two versions of the set of instruments, one which includes the US (IV) and the other which excludes the US ($IVnoUS$).

Discussion.—At this stage, it is important to point to a subtle yet non-trivial detail regarding the external IV strategy. As detailed in section 3.4.2, the goal of using external IVs is to eliminate endogeneity concerns caused by persistent unobserved factors that may be correlated with our proxies of interest ($proxies_{jct-1}$ in equation 6) or any of their lagged values. One pertinent example of an unobserved factor is domestic demand shocks, which could drive the way in which upstream and downstream industries engage in international trade. In the likely case that these demand shocks are also correlated with other key production inputs (k_{it}, l_{it}, and m_{it} in equation 6), internal IVs for these variables would now no longer be valid. For example, unobserved persistent domestic demand shocks are likely to be inter-temporally correlated with the firm's decision to fire/hire and/or invest in capital. As such, current and/or lagged values of the production inputs are no longer valid instruments. Otherwise stated, the use of external IVs for the proxies could invalidate the timing assumptions used to construct internal IVs for the production inputs that identify the production technology parameters (α_j in equation 6). If the case, one would *also* need to either find external IVs for these firm-level choice variables or impose additional assumptions that would guarantee the validity of internal IVs for the production inputs.

Finding external IVs for the production inputs, as far as we are concerned, is rather infeasible

[15]This includes: Australia; Brazil; Canada; India; Indonesia; Japan; Mexico; Russia; South Korea; Taiwan; Turkey; United States of America; and a rest of the world aggregate.

in a production function context; it is one of the main reasons for the advancement of both dynamic panel and proxy variable methods. Indeed, the quintessence of identification in this literature lies in the timing assumptions about when choices are made relative to the shocks that hit the firm (Ackerberg 2020). Therefore, to guarantee that the internal IVs for the production inputs are valid, we need to further assume that any persistent, unobserved factors are only correlated with $proxies_{jct-1}$ and not with k_{it}, l_{it}, and m_{it}.

3.5 Statistical Inference

To ensure precision in our standard error estimates, we implement a cluster (at the industry-country) bootstrap method.[16] This is important because the estimation procedure described above contains two steps. As such, closed form solutions for the variance-covariance matrices are not apparent. In addition, our regressors of interest are more aggregate than the regressand and thus take the same value for all observations within that level of aggregation, i.e. cluster. This induces a within-cluster error correlation that could lead to a downward bias of standard errors if not taken into account, and, resultantly, misleading estimator precision (Moulton 1986, 1990). Therefore, cluster bootstrap standard errors can be used for statistical inference similar to any other asymptotically valid standard errors. Online Appendix C provides a detailed description of the bootstrap procedure and the calculation of cluster bootstrap standard errors.

A subtle yet important point is that, compared to standard bootstrap methods, additional action needs to be taken given that our estimation procedure is an over-identified GMM. In this empirical setup, the population moment conditions do not hold exactly in the bootstrap sample. Consequently, the bootstrap and original sample versions of the test statistics would have different asymptotic distributions. For reliable bootstrap inference, we follow Hall and Horowitz (1996) in recentering the bootstrap version of the moment conditions relative to the original sample version. Online Appendix C.2 describes the implementation of the recentering.

Finally, to support the validity of the estimated model we conduct the following two tests. First, we test the validity of the additional moment conditions using a cluster bootstrap version of the Hansen-J statistic of overidentifying restrictions (Hansen 1982). Under the assumption that the baseline model is correctly specified, the null hypothesis is that the overidentifying moment conditions are valid. Second, we test for weak-instruments using a cluster bootstrap version of the underidentification test statistic established by Windmeijer (2021). Specifically, we test the null hypothesis that the model is underidentified, i.e. the instruments have insufficient explanatory power to predict the endogenous variable(s), and thus identify the model parameters. Overall, rejection of the overidentification test or non-rejection of the underidentification test would raise concerns about the model. Online Appendix B and C.3 explain in detail the construction of the test statistics and how to obtain their cluster bootstrap p-values, respectively.

[16]For bootstrap methods see: Efron (1979); Efron (1982); Horowitz (2001); and Davidson and MacKinnon (2004). For a complete review on cluster-robust inference see Cameron and Miller (2015).

4 Data

We construct a firm-level panel of manufacturing firms in 19 EU countries[17] from 2000 to 2014 from the Amadeus database by Bureau van Dijk Electronic Publishing (2018) (BvDEP). BvDEP regularly updates the information set in Amadeus and releases a monthly version which contains the latest information on ownership. Firms that exit the market are dropped fairly rapidly. For a complete set of financial and ownership information over time, we use a time series of (annual) releases to construct a consistent database. This allows us to build a dataset with nearly full financial and administrative information, i.e. balance sheet, profit and loss account activities, location, ownership, exit and entry. Merlevede et al. (2015) describe the construction and representativeness of the data at length.

We focus on the sample of active firms[18] that file unconsolidated accounts.[19] We retain firms which report their sales, tangible fixed assets, number of employees, material costs, ownership information and NACE 2-digit industry classification.[20] Following Merlevede et al. (2015), for better coverage and representativeness across EU countries, we keep firms with more than 20 employees on average. Table D.1 in Online Appendix D summarizes the remaining percentage coverage after applying each sample selection criterion. The selected sample captures the majority of reported sales and number of employees in the original sample, despite the sizeable drop in terms of the number of firms and firm-year observations. This is consistent with dropping a large mass of small-sized firms which by EU law have simplified reporting obligations in their annual financial statements (European Commission 2020a). Thus, our sample covers larger firms which constitute the bulk of manufacturing activity in terms of sales and employment (for an in depth discussion see Merlevede et al. 2015).

Next, we remove outliers using the BACON method proposed by Billor et al. (2000) to ensure that such observations do not drive overall results.[21] Lastly, to retain the same number of firms for both models in levels and first-differences, we keep firm-year observations observed for more than two consecutive years. Overall, this results in an unbalanced panel of 130,377 manufacturing firms and 1,223,955 observations across the 19 EU countries in our dataset for the period 2000-2014 (see Table 1).

Monetary variables are deflated using the appropriate country-specific NACE 2-digit output deflator from the EU KLEMS database (EU KLEMS 2017). Real output (Y) is sales deflated with

[17] Austria (AT); Belgium (BE); Bulgaria (BG); Croatia (HR); Czech Republic (CZ); Estonia (EE); Finland (FI); France (FR); Germany (DE); Hungary (HU); Italy (IT); Norway (NO); Poland (PL); Portugal (PT); Romania (RO); Slovakia (SK); Slovenia (SI); Spain (ES); and Sweden (SE).

[18] We exclude firms that are dissolved, in liquidation, inactive or in bankruptcy since their assets can genuinely go down to (almost) zero.

[19] Refers to accounts which do not integrate the statements of possible controlled subsidiaries or branches of the concerned company. This avoids double counting at the firm level but is uninformative about the plant dimension.

[20] Appendix Table A.1 provides an overview of the NACE Rev.2 2-digit industries included (Eurostat 2020).

[21] BACON is a multiple outlier detection method. The variables considered in the method are: the log of output, labour, capital, and material; and material's revenue share. We first trim at the industry and then manufacturing level. As in any outlier detection method, the threshold defining the outlying points is chosen by the researcher, therefore, we will provide a robustness check over the choice of a more lenient threshold.

producer price indices. Capital (K) is tangible fixed assets deflated by the average of the deflators of various NACE 2-digit industries (Javorcik 2004).[22] Real material inputs (M) is material inputs deflated by an intermediate input deflator constructed as a weighted average of output deflators, where country-industry-year specific weights are based on intermediate uses retrieved from IO-tables. Labour (L) is the number of employees. Finally, SUB^{dom} and SUB^{for} are dummy variables indicating whether the firm controls more than 10% of the shares of a domestic or foreign firm, respectively. SHH^{dom} and SHH^{for} are dummy variables indicating whether more than 10% of the firm's shares are owned by a domestic or foreign firm, respectively.

Table 1: Summary statistics

Firm-level	Obs.	Mean	St.Dev.	p25	p50	p75
Sales†	1,223,955	32,562	233,108	2,553	6,347	17,696
Tang. Fixed Assets†	1,223,955	7,337	73,537	320	1,105	3,687
Material†	1,223,955	19,630	180,711	1,009	2,954	9,248
Number of Employees	1,223,955	128	368	30	48	107
SUB^{dom}	1,223,955	.05	.22	0	0	0
SUB^{for}	1,223,955	.044	.21	0	0	0
SHH^{dom}	1,223,955	.22	.41	0	0	0
SHH^{for}	1,223,955	.099	.3	0	0	0
Industry-country-level						
$downIM$	5,265	0.352	0.227	0.161	0.326	0.512
$upEX$	5,265	0.209	0.120	0.124	0.184	0.266
$upIM$	5,265	0.210	0.069	0.159	0.202	0.248
$downEX$	5,265	0.217	0.106	0.134	0.207	0.289
IM	5,265	0.386	0.162	0.264	0.368	0.485
EX	5,265	0.459	0.263	0.241	0.445	0.671

Notes: † Monetary variables in thousands of Euro. Unbalanced panel of 130,377 firms in 19 manufacturing industries across 19 EU countries over the period 2000-2014.

For the measurement of proxies we use the WIOD November 2016 Release (henceforth WIOD), which provides a time series of World IO Tables (WIOT) for 43 countries worldwide and a table covering the rest of the world for the years 2000-2014.[23] WIOD uses the Statistical classification of products by activity (CPA) which contains 56 industries, 19 out of which are in the manufacturing sector.[24] A major advantage over other databases is that the WIOT varies over time and that information on goods imports does not rely on the standard proportionality assumption. Instead, a more flexible approach is followed whereby import proportions vary over end-use categories. This provides greater variability over time and intermediate input types. This extra level of detail is expected to unmask possible heterogeneity and provide

[22] Office machinery and computing (26); electrical machinery and apparatus (27); machinery and equipment (28); motor vehicles, trailers, and semi-trailers (29); and other transport equipment (30).

[23] See Timmer et al. (2015, 2016) for a detailed description of the construction of the tables.

[24] See Appendix Table A.1 for correspondence with the NACE Rev.2 2-digit industry classification.

better identification. The bottom panel of Table 1 presents summary statistics for the firm and industry-country-level variables in our sample.

5 Results

In this section, using the empirical methodology described above, we analyse the productivity effects of inter-industry importing and exporting in detail through: (a) presenting our baseline results; (b) conducting a battery of robustness checks; (c) establishing causality; and (d) uncovering heterogeneity that points to possible mechanisms in place. For ease of comparison, the first column of all results tables presents the baseline results.

5.1 TFP Effects from Inter-industry Importing and Exporting

In Table 2 we first introduce the results from our baseline specification (6) (column 1). We then test the robustness of our findings for: a number of alternative assumptions in the estimation procedure (columns 2-5) and the construction of the variables (columns 6-7).[25] Finally, Table 4 provides results from the alternative IV strategy to pin down the causal interpretation of our estimates. The key finding is that upstream exporting is the only robust channel for inter-industry effects of importing and exporting on firm performance. Otherwise stated, domestic sourcing from industries that also export intermediates leads to higher productivity levels for a given firm.

Baseline.—Column 1 of Table 2 shows the results for our main variables of interest when estimating specification (6).[26] Upstream exporting ($upEX_{jct-1}$) has a statistically significant effect on firm performance; a one standard deviation increase in upstream exporting enhances TFP by 0.75% in the short run and 2.58% in the long run.[27] While downstream exporting ($downEX_{jct-1}$) also has a statistically significant effect, it disappears or is weakly statistically significant under various subsequent robustness checks and the alternative IV strategy (Table 4).

The results for downstream importing ($downIM_{jct-1}$) and upstream importing ($upIM_{jct-1}$) reveal negative magnitudes which are—for the large part—not statistically significant. With regards to $downIM_{jct-1}$, while our results do not validate the same channel explored in Blalock and Veloso (2007) this is not concerning because we focus on a multi-country panel (versus one country only). Therefore, other mechanisms could be at play. For example, competition from abroad for domestic downstream clients might increase the costs of the focal firm's inputs

[25]We also test the robustness of our findings under heterogeneous firm sizes and additional checks. Results are presented in Online Appendix Table D.4.

[26]The estimation yields economically sensible estimates for the output elasticities of inputs and returns to scale (see Online Appendix Table D.2), which brings confidence over the reasonable performance of the baseline model.

[27]Using point estimates from baseline equation (6), the long-run effect is computed using the following formula: $\rho_{upEX}/(1-\rho_\omega) * sd(upEX_{jct-1}) * 100$, where ρ_{upEX} is the estimated short-run effect of lagged upstream exporting on TFP, ρ_ω is the estimated persistence of TFP, and $sd(upEX_{jct-1})$ is one standard deviation of the residuals generated after regressing the lagged values of upstream exporting on the time-varying fixed effects ($\phi_{fe,t}$) and firm fixed effects (ϕ_i), as specified in equation (6).

and hence reduce its profit margins. In turn, this would dampen its revenue productivity. This hypothesis is supported by the negative effect of downstream importing on Belgian firms' employment in Merlevede and Michel (2020). Following Defever et al. (2020), a similar reasoning could apply with regards to $upIM_{jct-1}$. Specifically, the authors find a negative effect in the case of domestic firms in industries which rely less intensively on wholesalers that import inputs from abroad.[28] However, given that the point estimates for these two effects are statistically insignificant, heavy weight should not be placed on their interpretation.

Table 2: TFP effects from inter-industry importing and exporting under baseline specification and robustness to alternative assumptions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline	Alternative Assumptions				Fix Weights in 2000	
		Fixed Effects		Labour	Imperfect	All	*downIM*
		ct	*jt*	Timing	Competition	Proxies	& *upEX*
ω_{it-1}	0.710***	0.756***	0.713***	0.677***	0.706***	0.714***	0.712***
	(0.022)	(0.014)	(0.017)	(0.021)	(0.020)	(0.022)	(0.022)
$downIM_{jct-1}$	-0.136	-0.120	-0.147	-0.141	-0.138	-0.137	-0.137
	(0.091)	(0.087)	(0.098)	(0.095)	(0.089)	(0.111)	(0.102)
$upEX_{jct-1}$	0.432***	0.265**	0.387**	0.433**	0.555***	0.315**	0.214**
	(0.166)	(0.121)	(0.163)	(0.181)	(0.181)	(0.129)	(0.109)
$upIM_{jct-1}$	-0.228	0.052	-0.269	-0.226	-0.299	-0.563**	0.083
	(0.207)	(0.205)	(0.211)	(0.222)	(0.216)	(0.260)	(0.195)
$downEX_{jct-1}$	0.267**	0.196*	0.244*	0.288**	0.244*	0.469**	0.204
	(0.134)	(0.106)	(0.126)	(0.146)	(0.132)	(0.206)	(0.135)
Observations	1,018,643	1,018,643	1,018,643	1,018,643	1,018,643	1,018,643	1,018,643

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. All regressions include: intra-industry importing and exporting; dummies for domestic and foreign ownership links; year, industry and country fixed effects; and industry and country linear time trends. Columns 2 and 3 also include country-year (ct) and industry-year (jt) fixed effects, respectively. Standard errors are computed using a cluster (at the industry-country) bootstrap with 99 replications and are reported in parentheses below point estimates.

Alternative Assumptions.—To exclude the possibility our results are being driven by any type of country-specific growth trends, we augment the baseline specification with country-time (ct) fixed effects (column 2). Coefficients for both upstream and downstream exporting are in line with the baseline estimates (column 1), but with their magnitudes reduced and their statistical significance weakened since the new fixed effects control for important sources of variation. Note, however, that various industries display different business cycle volatilities. For example, manufacturing of durable goods falls particularly severely during recessions. One might therefore be concerned that the effects we pick up reflect cyclical industrial variation (as both international trade and productivity are known to decline during recessions). To control for this, we instead augment the baseline specification with industry-time (jt) fixed effects (column

[28]To further validate this channel, we would need additional information on industries' wholesaler intensity.

3). We find that results are in line with those in column 1, but with their magnitudes reduced and their statistical significance weakened similar to column 2.[29]

Until now, we have performed estimations under the assumption that rigidities in the European labour market prevent the adjustment of labour within the (accounting) year. This translates to a one period lag between the choice of labour and its realisation in the production process (hence in the accounting data). Therefore, labour, as in the case of capital, is predetermined in period t. In column 4 we conduct an additional robustness test by instead assuming that the labour market is more flexible but still subject to adjustment costs. Under this assumption, labour is chosen during the realisation of productivity, i.e. between $t-1$ and t and thus correlated with the error term. Thus, in the second step of our estimation procedure, we now use l_{it-2} instead of l_{it-1} in the orthogonality conditions of the first differenced equation and Δl_{it-1} instead of Δl_{it} for the equation in levels. Under such alternative timing assumptions, results remain similar to column 1 with an expected increase in standard errors (Ackerberg 2020).

In column 5, we partly control for unobserved variation in output prices by introducing more structure under additional assumptions, following the methodology proposed by GNR. These assumptions include a time-varying iso-elastic CES demand system under monopolistic competition, similar to Klette and Griliches (1996).[30] This is expected to be insightful to the extent to which, on average, firms adjust their markups over time. However, for any deviation from this assumption we would need more detailed data, e.g. firm-level output prices, which is not currently available for such an extensive cross-country dataset. Therefore, as far as our approach accounts for price differences in the output market, the effects presented in column 5 remain similar to those in baseline in column 1, but with an increase (decrease) in magnitude and statistical significance in upstream exporting (downstream exporting).[31]

Construction of Variables.—To test the robustness of the weights used to build our proxies of interest, we fix weights to their values in year 2000 (the start year of the sample). This is done to eliminate potential distortions which may arise due to differences in the evolution of importing or exporting across time, countries, and industries. Conducting this check is important, since the weights used refer to domestic transactions. As such, they could potentially be endogenous if affected over time by the importing/exporting behaviour of the linked industries. In column 6 we fix the weights for all proxies to year 2000 while in column 7 we fix the weights for downstream importing and upstream exporting only. The latter step is taken because downstream importing and upstream exporting measures are more likely to be directly affected since the decision to source or supply domestically could be influenced by the intensity of importing or exporting,

[29] Recall that the variables of interest vary at the industry-country-year level. Under this set of fixed effects, the leftover identifying variation is limited. As a result, we might be asking too much from the data since we condition on potentially relevant dimensions of identifying variation. In addition, computationally both estimations are extremely time consuming and we thus avoid using them as our baseline.

[30] An exact description of the estimation procedure can be found in Appendix O5-4 of GNR. On top of the estimated effects of interest, this strategy identifies aggregate time-varying markups (see Online Appendix Figure D.3).

[31] Online Appendix Table D.2 confirms the downward bias in the output elasticities of inputs when assuming away imperfect competition in the output market (Klette and Griliches 1996).

respectively. As in column 1, upstream exporting does not seem to be driven by the choice of a specific weighting scheme, however, downstream exporting is now insignificant in column 7.

Given the relatively high level of aggregation, we find a moderate to high correlation between the proxies (Table 3). To avoid potential issues stemming from multicollinearity, we re-estimate the baseline specification by introducing each inter-industry proxy separately. We do this both when intra-industry proxies are included and excluded. Results are presented in columns 2-10 from Table D.3 in Online Appendix D. Overall, results remain similar to the baseline model in column 1, and thus do not appear to be driven by the presence of multicollinearity in the proxies.

Table 3: Correlation of proxies

	downIM	*upEX*	*upIM*	*downEX*	*IM*	*EX*
downIM	1.000					
upEX	0.661	1.000				
upIM	0.570	0.677	1.000			
downEX	0.531	0.491	0.641	1.000		
IM	0.605	0.526	0.657	0.594	1.000	
EX	0.735	0.471	0.500	0.484	0.645	1.000
downIM	1.000					
upEX	0.429	1.000				
upIM	0.474	0.477	1.000			
downEX	0.407	0.241	0.375	1.000		
IM	0.414	0.365	0.409	0.231	1.000	
EX	0.613	0.220	0.407	0.268	0.437	1.000

Notes: The top panel reports correlations between proxies. The bottom panel reports correlations between the residuals generated after regressing each proxy on the fixed effects and trends (ϕ's) in baseline specification (6).

For completeness, we also consider two alternative, but less precise, ways of constructing the proxies. The first includes the intra-industry supply of intermediates from the diagonal elements of the IO tables, i.e. the measures in (1)-(4) also incorporate elements when $d = j$ and $u = j$ in the summation. The second considers all supply chain connections, as opposed to only the direct domestic supply chain connections $j - d$ and $u - j$ for *downIM* and *upEX*, respectively. Otherwise stated, measures (1) and (2) include the 'overall' import (Φ_{dct}) and export (Ψ_{uct}) intensities instead of those based only on the product links with downstream (Φ_{jdct}) and upstream (Ψ_{juct}) industries, respectively (see Figure 1 and discussion in Section 2).

Results point to upstream exporting as the only robust channel even under these less informative ways of constructing the measures (see columns 2 and 3 in Table D.4 in Online Appendix D). In terms of magnitudes, the estimate of *upEX* for the first set of measures is similar to the baseline. This is expected, since the additional variation from the diagonal elements in the measures is already accounted for from the intra-industry proxies included in the estimation. For the second set of controls, the estimate is almost twice as large as the baseline. This result suggests that additional productivity spillovers could arise when intermediates are exported by upstream industries but not supplied to domestic downstream industries.

Firm Size.—In columns 4 and 5 from Online Appendix Table D.4 we split our sample into small and medium-large firms, respectively.[32] Grouping medium and large firms allows to split the sample in two, since 99% of European firms are small and medium-sized (European Commission 2020b). Further, we choose this cut-off since finance and EU support programmes are targeted specifically at small size firms. Overall, this specific cut-off ensures that we observe both small and medium-large firms in all industries considered. We find that our baseline result for upstream exporting is highly significant for small firms, while it becomes weakly significant for large firms.[33] To formally test the differences across groups, we next estimate the empirical model by interacting the variables of interest with a dummy equal to one when a firm is small and zero otherwise. Results in column 6 of Online Table D.4 reveal that the difference in TFP effects from upstream exporting between small and medium-large firms is not statistically significant. In other words, small firms do not seem to be affected differentially by upstream exporting. This may reflect the fact that firm characteristics other than size could impact the extent to which firms are exposed to the effects of internationalisation through the domestic supply chain.

Additional Robustness.—To further support the validity of our results we proceed with a battery of additional robustness checks which are presented in columns 7-11 of Table D.4 in Online Appendix D. In column 7, we relax the threshold for dropping outliers from the 30th to the 15th percentile of the distribution of nominated outliers using the BACON procedure.[34] In column 8, we bootstrap the baseline specification for 499 replications, while in columns 9-11 we bootstrap the standard errors at the industry (j), country (c), and firm (i) cluster, respectively.[35] In all of the above cases, results remain robust. Finally, to check the sensitivity of standard errors in the presence of too few clusters, we also examine the distribution of bootstrapped point estimates (see Online Appendix Figure D.4). The unimodality we find suggests that standard errors are robust to the presence of potential outlier clusters (Cameron and Miller 2015).

Alternative IV Strategy.—Turning to the alternative IV strategy, we now discuss results when replacing internal IVs from our baseline specification with external IVs. Recall that we experiment with two different versions, including and excluding the US from the construction of the instruments because of the EU's strong links with this economy in particular. Results are presented in columns 2 and 3 of Table 4, respectively. They point to the causal interpretation of upstream exporting. Estimates of upstream exporting are very similar to one another irrespective

[32]On average, small firms report less than 50 employees and up to ten million Euro turnover. Medium-large firms are above this cut-off (European Commission 2020b).

[33]Splitting the sample implies that the production technology $f_j(\cdot)$ in equation (6) now differs across size groups.

[34]This results in an additional 5,581 firms and 54,399 firm-year observations.

[35]These results also help compute standard errors in the case that clustering arises at multiple levels simultaneously (Cameron et al. 2011). However, in this application, multi-way clustered standard-errors are expected to remain similar since, for one-way clustering with the fewest number of clusters, standard errors are only modestly different from the baseline model (Cameron and Miller 2015). Furthermore, there is no clear reason why additional error correlation in other dimensions needs to be high and thus bias downwards the standard errors, especially when we also control for fixed effects in those dimensions.

of the inclusion of the US in the external instrument, and slightly larger in magnitude than those in the baseline model with the internal IV strategy (column 1).

The difference in magnitudes between column 1 and columns 2-3 suggests, if anything, a correlation between our variables of interest and the error term which induces a downward bias in our baseline estimates. To illustrate the potential issue, consider an expansionary shock to the domestic industry j. Because of its increased need for domestic intermediates, firms in upstream industries u now need to rely relatively less on selling their inputs abroad. This generates a negative correlation between domestic shocks in industry j and upstream exporting.[36]

Table 4: TFP effects from inter-industry importing and exporting with alternative IV strategy

	(1) Baseline	(2) *IV*	(3) *IVnoUS*
$downIM_{jct-1}$	-0.136	-0.167	-0.057
	(0.091)	(0.111)	(0.106)
$upEX_{jct-1}$	0.432***	0.450**	0.463**
	(0.166)	(0.209)	(0.201)
$upIM_{jct-1}$	-0.228	-0.367	-0.450
	(0.207)	(0.273)	(0.304)
$downEX_{jct-1}$	0.267**	0.000	-0.162
	(0.134)	(0.154)	(0.141)
Overidentification p-value	0.475	0.141	0.081
Underidentification p-value for...			
$downIM_{jct-1}$	0.000	0.000	0.000
$upEX_{jct-1}$	0.000	0.000	0.000
$upIM_{jct-1}$	0.000	0.000	0.000
$downEX_{jct-1}$	0.000	0.000	0.000
Observations	1,018,643	1,018,643	1,018,643

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. All regressions include: the persistence term; intra-industry importing and exporting; dummies for domestic and foreign ownership links; year, industry and country fixed effects; and industry and country linear time trends. Standard errors are computed using a cluster (at the industry-country) bootstrap with 99 replications and are reported in parentheses below point estimates. Overidentification p-value is the bootstrap p-value for the overidentification Hansen-*J* test under the null of valid overidentifying restrictions. Underidentification p-value is the bootstrap p-value on the per endogenous explanatory variable underidentification test under the null that each endogenous variable is poorly predicted by the instruments (see Section 3.5 and Online Appendix B).

The lower panel of Table 4 provides the bootstrap p-values of the overidentification Hansen-*J*

[36]Given the various levels at which the shocks could arise, there are numerous plausible ways to illustrate the direction of this bias. However, the bias crucially depends on the sign of both the correlation between the unobserved shocks with the variables of interest and the impact of those unobserved shocks on productivity. See Basu (2020) for a detailed discussion on this topic and the difficulties in determining the direction of a bias.

test statistic. We fail to reject the null hypothesis of valid overidentifying restrictions. Similarly, we reject the null hypothesis of underidentification, as seen by the bootstrap p-values of the underidentification test on each endogenous variable (rows) across all model specifications (columns 1-3). These results bring further confidence about the model specification and its identification strength under both the internal and external IV strategies—keeping in mind the caveats of those test statistics (Roodman 2009) and the relevant considerations in interpreting the magnitude of p-values (Kiviet 2020).

Overall, Table 4 shows that results from the external IVs are in line with those in our baseline model. Nonetheless, as discussed in Section 3.4.2, the choice of external IVs requires strong assumptions to guarantee that the internal IVs for the production inputs remain valid. Because of this, and coupled with the fact that we get similar point estimates, we base the rest of our analysis on estimating variants of column 1, Table 2. Therefore, the purpose of the following exercises is to explore patterns in the data that would provide suggestive evidence of heterogeneous effects and possible mechanisms in place, rather than definitively establishing causality.

5.2 Heterogeneity

Our results for firm size raise the question as to whether the productivity effects from upstream exporting are applicable to firms that do not (or are less likely to) directly participate in domestic supply chain activity. We explore this heterogeneity in Table 5 by testing for differential effects in our baseline setup. Specifically, we interact our proxies with a dummy variable D that is equal to one when a firm belongs to the group defined at the top of each column, and zero otherwise (columns 2-7).[37] In what follows, we discuss each group in turn.

Domestic Supply Chain Participation.—Purely foreign owned firms are more likely to be directly involved in an international supply chain relative to local firms with no foreign ownership links. The latter are more likely to depend on domestic clients and suppliers. Thus, we may expect firms without any foreign ownership links to be more prone to inter-industry productivity effects. Column 2 confirms this intuition. As seen by the interaction term ($D * upEX_{jct-1}$), a significantly larger TFP effect is detected from upstream exporting for the group of firms that have no foreign (but possibly domestic) ownership links. This suggests that the domestic supply chain may act as a vehicle for internationalisation effects through upstream exporting.

Supply Chain Position.—Fally (2012) finds a large shift of value-added towards final stages of production, i.e. relatively downstream. He further shows that richer countries have a comparative advantage in goods that involve fewer production stages and goods that are closer to final demand.[38] This is also in line with Antràs et al. (2012), who show that a better rule of law, strong financial development, and relatively high skill intensity are correlated with a larger propensity to export in relatively more downstream industries. This translates to the fact that

[37] Table D.5 in Online Appendix D repeats the analysis but with sample splits instead of dummy interactions.

[38] The latter stylised fact is consistent with the theoretical predictions of Costinot et al. (2013).

relatively downstream industries both sell domestically and export output for final use more intensively. Therefore, relatively less output is expected to be available for domestic supply in relatively downstream industries. Conversely, relatively downstream industries will have more upstream opportunities, i.e. relatively more output sourced from upstream, compared to relatively upstream industries.

Table 5: Heterogeneity in TFP effects from inter-industry importing and exporting

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		D=1 if a firm belongs to the group described below and 0 otherwise					
	Baseline	Domestic Links	Relatively Downstream	Relatively Low-tech	Labour Intensive	Downstream & Low-tech	Downstream & Labour-int
$downIM_{jct-1}$	-0.136	-0.116*	0.042	0.146	0.053	0.149	0.164
	(0.091)	(0.063)	(0.146)	(0.097)	(0.097)	(0.098)	(0.102)
$upEX_{jct-1}$	0.432***	0.198*	-0.471**	0.044	0.008	0.072	0.084
	(0.166)	(0.108)	(0.239)	(0.125)	(0.147)	(0.134)	(0.139)
$upIM_{jct-1}$	-0.228	-0.234	0.651**	0.002	0.092	-0.035	-0.061
	(0.207)	(0.191)	(0.253)	(0.225)	(0.274)	(0.209)	(0.210)
$downEX_{jct-1}$	0.267**	0.150	0.300	0.125	0.007	0.215	0.194
	(0.134)	(0.111)	(0.192)	(0.140)	(0.176)	(0.133)	(0.132)
$D*downIM_{jct-1}$		-0.017	-0.304*	-0.451***	-0.428***	-0.471**	-0.457**
		(0.049)	(0.163)	(0.153)	(0.138)	(0.185)	(0.181)
$D*upEX_{jct-1}$		0.222***	1.138***	1.547***	1.176***	1.630***	1.565***
		(0.082)	(0.298)	(0.318)	(0.295)	(0.387)	(0.371)
$D*upIM_{jct-1}$		-0.058	-1.091**	-0.968*	-1.057**	-1.092*	-1.024*
		(0.096)	(0.480)	(0.530)	(0.466)	(0.628)	(0.613)
$D*downEX_{jct-1}$		0.051	-0.127	0.103	0.345	0.015	-0.036
		(0.055)	(0.262)	(0.245)	(0.241)	(0.300)	(0.298)
Observations	1,018,643	1,018,643	1,018,643	1,018,643	1,018,643	1,018,643	1,018,643

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. All regressions include: the persistence term; intra-industry importing and exporting; dummies for domestic and foreign ownership links; year, industry and country fixed effects; and industry and country linear time trends. D is a dummy variable equal to one for firms in the group described in each column and zero otherwise. D is time invariant and thus not separately identified from the fixed effects, except in column 2 where it is time varying but not reported for space considerations. Columns 2-7 include interactions of intra-industry importing and exporting with D. Standard errors are computed using a cluster (at the industry-country) bootstrap with 99 replications and are reported in parentheses below point estimates.

As a result, more domestic upstream relationships will be generated in relatively downstream industries. This leads us to expect a larger potential for upstream exporting in relatively downstream industries.[39] To account for possible heterogeneity in the absorption of TFP effects from inter-industry linkages, we generate an industry-level measure of relative supply chain position as in Fally (2012) and Antràs et al. (2012) using WIOT. This measure of upstreamness

[39]Theoretical predictions of Antràs and Chor (2013) show that the incentive to integrate suppliers varies systematically with the relative position at which the supplier enters the production line. The nature of the relationship between integration and downstreamness depends crucially on the elasticity of demand faced by the final good producer and the degree of complementarity between inputs in production. However, for the case of the 19 European countries in our analysis, we are not aware of any such elasticities. Therefore, we do not have any expectations about the exact direction of results. Nonetheless, we do expect that relatively upstream or downstream firms are differentiated in how they absorb the inter-industry effects depending on their integration intensities.

gives the average 'distance' of each industry from final use. We rank industries as relatively downstream or upstream based on the median value of the distribution of this measure.[40]

Column 3 of Table 5 presents results when interacting the proxies with a dummy that is equal to one for firms in relatively downstream industries and zero otherwise, i.e. relatively upstream. We confirm that firms in relatively downstream industries experience significantly larger productivity effects from upstream exporting which appear as non-existent for relatively upstream industries.

High-tech versus Low-tech.—Fally (2012) also finds that R&D intensive industries have become relatively less fragmented over time. This complements the work of Acemoglu et al. (2007, 2010) who find that innovative industries rely less intensively on outsourcing. Therefore, we might expect relationship-specific learning that depends on links between firms through their exchange of intermediate inputs to be less limited in low-tech industries where such relationships are more prevalent. On the basis of information on the technological intensity of industries (available in Eurostat 2018), we construct a dummy variable equal to one when firms are in relatively low-tech industries and zero otherwise.[41] In column 4 of Table 5 we interact the proxies and dummy variable. We find that firms in low-tech industries experience significantly larger productivity effects from upstream exporting.[42]

Labour versus Capital Intensive.—Antràs (2003) shows that there is a higher propensity for integration in capital intensive industries. Therefore, labour intensive industries are expected to rely more on outsourcing and consequently experience stronger productivity effects from inter-industry importing and exporting. In column 5 of Table 5 we therefore interact the proxies with a dummy variable equal to one when a firm is in a relatively labour intensive industry and zero otherwise.[43] As expected, we find that on average firms in relatively labour intensive industries experience statistically larger productivity effects from upstream exporting.

Combinations.—For additional heterogeneity, we lastly create new groups based on combinations of the previous categories. In columns 6 and 7 of Table 5, the dummy variable is now the intersection of downstream & low-tech industries and downstream & labour intensive industries,

[40] See Table D.7 in Online Appendix D. Relatively downstream includes industries with CPA: 12; 22; 21; 17; 20; 6; 5; 19; and 18. Relatively upstream includes industries with CPA: 10; 23; 11; 13; 16; 7; 14; 8; 15; and 9. As in Antràs et al. (2012), we observe that primary and resource-extracting industries tend to be relatively upstream.

[41] Eurostat (2018) groups manufacturing activities to 'high-technology,' 'medium high-technology,' 'medium low-technology' and 'low-technology' based on the R&D expenditure/value added of industries. We define as relatively low-tech the industries in the 'low-technology' group with CPA: 5; 6; 7; 8; 9; and 22. The rest of the manufacturing industries are considered as relatively high-tech.

[42] One could argue that these effects also depend on the country's level of economic development. Therefore, in columns 2 and 3 of Table D.6 in Online Appendix D we repeat the analysis by splitting the sample into Central Eastern and Western European Countries, respectively. We find that only Central Eastern European firms in low-tech industries appear to benefit from upstream exporting. These findings point to learning as a plausible mechanism for explaining our results, which we discuss in detail in the next section.

[43] Labour intensity is computed using industry-country-year-level data on capital compensation and number of employees from the Socio Economic Accounts of WIOD (Release 2016). Labour-int industries are defined based on values smaller than the median of the distribution of average capital to labour ratios for each industry across countries and over the period considered. Labour-int includes industries with CPA: 5; 6; 7; 9; 15; 16; 21; 22; and 23, and capital-int includes the rest of the manufacturing industries.

respectively. As before, we expect these groups to generate relatively more intermediate input based relationships than their reference groups ($D = 0$). In line with the previous results, we find significantly larger productivity effects from upstream exporting for these groups.

5.3 Exploration of the Underlying Mechanism

So far, we have established that the robust productivity effect from upstream exporting is prevalent in firms and industries that are less likely to be directly involved in international trade, i.e. at lower parts of the productivity distribution. This coincides with the results of several other studies, summarised in Bernard et al. (2012). However, this result alone does not inform us of the process through which these firms upgrade their efficiency. As discussed in section 2, learning is a key candidate mechanism by which to explain our results. In order to capture the potential presence of this mechanism, in Table 6 we exploit yet additional heterogeneity, stemming from the level of development of both the countries in our sample and the trading partner countries in WIOD. As discussed below, results support learning as a key mechanism to explain the robust productivity premia from upstream exporting—which are dominant in countries and/or industries at lower levels of development.

Central Eastern European Countries versus Western European Countries.—If learning mechanisms are in place, we would expect firms in less developed countries to benefit more from upstream exporting, since the scope for improving production and quality is expected to be higher (Fernandes and Isgut 2015). This reasoning is also in line with findings whereby domestic firms in transition economies benefit from FDI spillovers through the supply chain (Havranek and Irsova 2011). In column 2 of Table 6 we therefore interact the proxies with a dummy variable equal to one when firms are located in a Central Eastern European Country (CEEC) and zero otherwise, i.e. in a Western European Country (WEC).[44] Results suggest that firms in CEEC benefit the most from upstream exporting.

Developed versus Developing Trading Partners.—De Loecker (2007) finds productivity premia for Slovenian firms (also in our sample) exporting to high-income countries. This supports the learning by exporting hypothesis. Therefore, if there is scope for learning from foreign markets through upstream domestic suppliers, we expect to find larger efficiency gains when suppliers ship products to relatively more advanced economies. However, these gains from learning should be apparent only in countries that are further away from the technological frontier (and where there is more room for catching up). Conversely, if there are mechanisms other than learning at play (e.g. self-selection, cost advantage, competition, etc.) we expect to find productivity effects from upstream exporting to developing regions. Therefore, in columns 3a and 3b of Table 6 we focus on the CEEC sample and, based on the trading partner countries in the country-by-industry entries of the WIOD, we split the proxies based on the trading partner's

[44] CEEC cover: Bulgaria; Croatia; Czech Republic; Estonia; Hungary; Poland; Romania; Slovakia; and Slovenia. WEC cover: Austria; Belgium; Finland; France; Germany; Italy; Norway; Portugal; Spain; and Sweden.

level of development.[45] Consistent with learning mechanisms, we find significant productivity premia in CEEC only from upstream exporting to developed countries.

Table 6: TFP effects from inter-industry importing and exporting with sample and proxy splits

	(1)	(2)	(3a)	(3b)	(4a)	(4b)
			Firms in CEEC		Firms in WEC	
	Baseline	D=1 if in CEEC	Proxies when trade partner: Developed	Proxies when trade partner: Developing	Proxies when trade partner: Developed	Proxies when trade partner: Developing
$downIM_{jct-1}$	-0.136	-0.072	0.059	-0.259	0.065	-0.374***
	(0.091)	(0.089)	(0.179)	(0.181)	(0.119)	(0.113)
$upEX_{jct-1}$	0.432***	-0.003	0.845**	0.275	-0.064	-0.080
	(0.166)	(0.160)	(0.407)	(0.283)	(0.335)	(0.301)
$upIM_{jct-1}$	-0.228	-0.275	0.483	-0.275	0.545	0.865***
	(0.207)	(0.360)	(0.738)	(0.277)	(0.472)	(0.295)
$downEX_{jct-1}$	0.267**	0.250	0.395	0.083	0.054	0.947***
	(0.134)	(0.153)	(0.257)	(0.284)	(0.297)	(0.367)
$D * downIM_{jct-1}$		-0.054				
		(0.148)				
$D * upEX_{jct-1}$		0.707***				
		(0.271)				
$D * upIM_{jct-1}$		0.661				
		(0.537)				
$D * downEX_{jct-1}$		0.062				
		(0.141)				
Observations	1,018,643	1,018,643	277,003		741,640	

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. All regressions include: the persistence term; intra-industry importing and exporting; dummies for domestic and foreign ownership links; year, industry and country fixed effects; and industry and country linear time trends. D is a dummy variable equal to one for firms in CEEC and zero otherwise. D is time invariant and thus not separately identified from the fixed effects. Column 2 also includes the interactions of intra-industry importing and exporting with D. Estimates in columns $3a(4a)$ and $3b(4b)$ are from the same estimation and represent the split of the proxies based on the level of development of the trading partner, i.e. developed and developing, respectively. For those cases we split the intra-industry proxies. Standard errors are computed using a cluster (at the industry-country) bootstrap with 99 replications and are reported in parentheses below point estimates.

Finally, in columns 4a and 4b we repeat our analysis for firms in the WEC sample. Results suggest that firms located in developed countries do not benefit from spillover effects arising from upstream exporting to developing countries. This finding is consistent with a small scope for firms that are based in developed countries to learn by exporting when those exports are destined to developing countries. Interestingly, in column 4b, we find statistically significant effects from

[45] Technological development is based on the countries' gross domestic spending on R&D as a percentage of GDP (OECD 2019). Developed countries include: Australia; Austria; Belgium; Canada; Denmark; Finland; France; Germany; Great Britain; Japan; the Netherlands; Norway; South Korea; Sweden; Switzerland; Taiwan; and United States of America. The rest of the countries in WIOD are defined as technologically less developed. (UNCTAD 2019).

downstream importing, upstream importing, and downstream exporting to developing countries. These findings might indicate that in WEC mechanisms other than learning might be at play for those effects, e.g. cost advantage, competition, reallocation, among others (see Shu and Steinwender 2019; Defever et al. 2020). Thus, one avenue for future research could be to explore these mechanisms for firms in developed countries in particular. Overall, however, the main results in this section support learning as a key mechanism for the robust productivity premia generated by upstream exporting.

6 Conclusion

A large literature has examined the relationship between export and import behaviour on productivity. Yet, despite substantial research on the direct productivity effects of firms' or industries' trading behaviour, spillovers from internationalisation through the domestic supply chain have received less attention. In this paper, we examine whether the importing and exporting activities of a firm's suppliers and/or customers along their domestic supply chains can affect the firm's productivity, even if the firm itself may not be directly engaged in international trade. We provide novel evidence that upstream exporting, i.e. sourcing from upstream industries which also export intermediates, leads to productivity gains for the domestic client. Our analysis suggests that a one standard deviation increase in upstream exporting is associated with a productivity increase of 0.75% in the short run and 2.58% in the long run.

To pin down causality of our headline result, we implement an external IV strategy that exploits two identification ideas common in the literature. The first follows the suggestions of Autor et al. (2013) and Dauth et al. (2014) who draw on variation from trade flows of other high-income countries as instruments for domestic import and export exposure. The second exploits the arrival of China in the world market as a quasi-natural event. In both cases, identification comes from shocks abroad that drive the internationalisation behaviour of industries linked through the domestic supply chain, while excluding domestic shocks in the industry of the firm considered. Implementing this strategy reinforces our main results and supports the causal interpretation of our findings.

Further analysis suggests that the effects are stronger for firms and industries that are less likely to be directly internationally involved. They also appear to be stronger for firms with non-foreign ownership status and in relatively downstream, low-tech, or labour intensive industries (or combinations thereof). Moreover, our findings support that learning is a mechanism through which these productivity spillover effects occur. In line with the learning by exporting literature, we find the effect to be present only for firms in developing countries and/or when the trading partners are technologically advanced. Overall, firms that are further away from the technological frontier can use the domestic supply chain to access the benefits of internationalisation via learning practices.

Given the empirical nature of our analysis, we see rich potential for our findings to motivate theoretical models on firm heterogeneity, supply chains and trade. Such extensions would lend nicely to counterfactual predictions about patterns of trade, production and productivity from changes in policies related to internationalisation.

References

Acemoglu, D., P. Aghion, C. Lelarge, J. Van Reenen, and F. Zilibotti (2007). Technology, information, and the decentralization of the firm. *Quarterly Journal of Economics 122*(4), 1759–1799.

Acemoglu, D., R. Griffith, P. Aghion, and F. Zilibotti (2010). Vertical integration and technology: Theory and evidence. *Journal of the European Economic Association 8*(5), 989–1033.

Ackerberg, D. (2020). Timing assumptions and efficiency: Empirical evidence in a production function context. Working paper - unpublished.

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

Amiti, M. and J. Konings (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from Indonesia. *American Economic Review 97*(5), 1611–1638.

Antràs, P. (2003). Firms, contracts, and trade structure. *Quarterly Journal of Economics 118*(4), 1375–1418.

Antràs, P. and D. Chor (2013). Organizing the global value chain. *Econometrica 81*(6), 2127–2204.

Antràs, P., D. Chor, T. Fally, and R. Hillberry (2012). Measuring the upstreamness of production and trade flows. *American Economic Review: Papers & Proceedings 102*(3), 412–416.

Antràs, P., T. C. Fort, and F. Tintelnot (2017). The margins of global sourcing: Theory and evidence from US firms. *American Economic Review 107*(9), 2514–64.

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies 58*(2), 277–297.

Arellano, M. and O. Bover (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics 68*(1), 29–51.

Arrow, K. J. (1962). The economic implications of learning by doing. *Review of Economic Studies 29*(3), 155–173.

Autor, D. H., D. Dorn, and G. H. Hanson (2013). The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review 103*(6), 2121–68.

Autor, D. H., D. Dorn, and G. H. Hanson (2016). The China shock: Learning from labor-market adjustment to large changes in trade. *Annual Review of Economics 8*(1), 205–240.

Aw, B. Y., M. J. Roberts, and D. Y. Xu (2008). R&D investments, exporting, and the evolution of firm productivity. *American Economic Review 98*(2), 451–56.

Bas, M. and V. Strauss-Kahn (2015). Input-trade liberalization, export prices and quality upgrading. *Journal of International Economics 95*(2), 250–262.

Basu, D. (2020). Bias of OLS estimators due to exclusion of relevant variables and inclusion of irrelevant variables. *Oxford Bulletin of Economics and Statistics 82*(1), 209–234.

Bernard, A. B., B. J. Jensen, S. J. Redding, and P. K. Schott (2012). The empirics of firm heterogeneity and international trade. *Annual Review of Economics 4*(1), 283–313.

Billor, N., A. S. Hadi, and P. F. Velleman (2000). BACON: Blocked Adaptive Computationally Efficient Outlier Nominators. *Computational Statistics & Data Analysis 34*(3), 279–298.

Blalock, G. and P. J. Gertler (2008). Welfare gains from foreign direct investment through technology transfer to local suppliers. *Journal of International Economics 74*(2), 402–421.

Blalock, G. and F. M. Veloso (2007). Imports, productivity growth, and supply chain learning. *World Development 35*(7), 1134–1151.

Blundell, R. and S. Bond (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics 87*(1), 115–143.

Blundell, R. and S. Bond (2000). GMM estimation with persistent panel data: an application to production functions. *Econometric Reviews 19*(3), 321–340.

Bond, S. and M. Söderbom (2005). Adjustment costs and the identification of Cobb Douglas production functions. Technical report, IFS Working Papers, Institute for Fiscal Studies (IFS).

Bureau van Dijk Electronic Publishing (2018). Amadeus database. URL: `http://www.bvdinfo.com/en-us/our-products/company-information/international/amadeus`.

Bustos, P. (2011). Trade liberalization, exports, and technology upgrading: Evidence on the impact of MERCOSUR on Argentinian firms. *American Economic Review 101*(1), 304–40.

Cameron, A. C. and D. L. Miller (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources 50*(2), 317–372.

Cameron, C. A., J. B. Gelbach, and D. L. Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics 29*(2), 238–249.

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

Coe, D. T. and E. Helpman (1995). International R&D spillovers. *European Economic Review 39*(5), 859–887.

Connolly, M. (2003). The dual nature of trade: Measuring its impact on imitation and growth. *Journal of Development Economics 72*(1), 31–55.

Costinot, A., J. Vogel, and S. Wang (2013). An elementary theory of global supply chains. *Review of Economic Studies 80*(1), 109–144.

Dauth, W., S. Findeisen, and J. Suedekum (2014). The rise of the East and the far East: German labor markets and trade integration. *Journal of the European Economic Association 12*(6), 1643–1675.

Davidson, R. and J. G. MacKinnon (2004). *Econometric theory and methods*, Volume 5. Oxford University Press New York.

De Loecker, J. (2007). Do exports generate higher productivity? Evidence from Slovenia. *Journal of International Economics 73*(1), 69–98.

De Loecker, J. (2013). Detecting learning by exporting. *American Economic Journal: Microeconomics 5*(3), 1–21.

Defever, F., M. Imbruno, and R. Kneller (2020). Trade liberalization, input intermediaries and firm productivity: Evidence from China. *Journal of International Economics*, 103329.

Doraszelski, U. and J. Jaumandreu (2013). R&D and productivity: Estimating endogenous productivity. *Review of Economic Studies*.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics 7*(1), 1–26.

Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. Society for Industrial and Applied Mathematics.

EU KLEMS (2017). EU KLEMS growth and productivity accounts: Statistical module, ESA 2010 and ISIC Rev. 4 industry classification, September 2017 release. URL: `http://www.euklems.net/`.

European Commission (Accessed on 11 July 2020a). EU rules on financial information disclosed by companies. URL: `https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/financial-reporting_en`.

European Commission (Accessed on 11July 2020b). What is an SME? URL: `http://ec.europa.eu/growth/smes/business-friendly-environment/sme-definition_en`.

Eurostat (Accessed on 10 July 2018). Statistics on high-tech industry and knowledge intensive services. URL: `http://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:High-tech`.

Eurostat (Accessed on 11 June 2020). NACE Rev. 2 - statistical classification of economic activities. URL: `https://ec.europa.eu/eurostat/web/nace-rev2`.

Fally, T. (2012). Production staging: Measurement and facts. *Boulder, Colorado, University of Colorado Boulder*, 155–168.

Fernandes, A. M. and A. E. Isgut (2015). Learning-by-exporting effects: Are they for real? *Emerging Markets Finance and Trade 51*(1), 65–89.

Gandhi, A., S. Navarro, and D. A. Rivers (2020). On the identification of gross output production functions. *Journal of Political Economy 128*(8), 2973–3016.

Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova (2010). Imported intermediate inputs and domestic product growth: Evidence from India. *Quarterly Journal of Economics 125*(4), 1727–1767.

Görg, H., A. Hanley, and E. Strobl (2008). Productivity effects of international outsourcing: Evidence from plant-level data. *Canadian Journal of Economics/Revue canadienne d'Economique 41*(2), 670–688.

Griliches, Z. and J. Mairesse (1999). Production functions: The search for identification. In S. Strøm (Ed.), *Econometrics and Economic Theory in the 20th Century*, pp. 169–203. Cambridge University Press. Cambridge Books Online.

Grossman, G. M. and E. Helpman (1991). Quality ladders in the theory of growth. *Review of Economic Studies 58*(1), 43–61.

Grossman, G. M. and E. Helpman (1995). Chapter 25 technology and trade. Volume 3 of *Handbook of International Economics*, pp. 1279–1337. Elsevier.

Hall, P. and J. L. Horowitz (1996). Bootstrap critical values for tests based on generalized-method-of-moments estimators. *Econometrica 64*(4), 891–916.

Halpern, L., M. Koren, and A. Szeidl (2015). Imported inputs and productivity. *American Economic Review 105*(12), 3660–3703.

Hansen, L. P. (1982). Large sample properties of Generalized Method of Moments estimators. *Econometrica 50*(4), 1029–1054.

Havranek, T. and Z. Irsova (2011). Estimating vertical spillovers from FDI: Why results vary and what the true effect is. *Journal of International Economics 85*(2), 234–244.

Horowitz, J. L. (2001). Chapter 52 - The bootstrap. Volume 5 of *Handbook of Econometrics*, pp. 3159–3228. Elsevier.

Javorcik, B. S. (2004). Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages. *American Economic Review 94*(3), 605–627.

Jovanovic, B. and Y. Nyarko (1996). Learning by doing and the choice of technology. *Econometrica 64*(6), 1299–1310.

Kasahara, H. and J. Rodrigue (2008). Does the use of imported intermediates increase productivity? Plant-level evidence. *Journal of Development Economics 87*(1), 106–118.

Kellogg, R. (2011). Learning by drilling: Interfirm learning and relationship persistence in the Texas oilpatch. *Quarterly Journal of Economics 126*(4), 1961–2004.

Kiviet, J. F. (2020). Microeconometric dynamic panel data methods: Model specification and selection issues. *Econometrics and Statistics 13*, 16 – 45.

Klette, T. J. and Z. Griliches (1996). The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics 11*(4), 343–361.

Levinsohn, J. and A. Petrin (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies 70*(2), 317–341.

Lileeva, A. and D. Trefler (2010). Improved access to foreign markets raises plant-level productivity…for some plants. *Quarterly Journal of Economics 125*(3), 1051–1099.

Markusen, J. R. (1989). Trade in producer services and in other specialized intermediate inputs. *American Economic Review 79*(1), 85–95.

Marschak, J. and W. H. Andrews (1944). Random simultaneous equations and the theory of production. *Econometrica 12*(3/4), 143–205.

Mendershausen, H. (1938). On the significance of professor Douglas' production function. *Econometrica 6*(2), 143–153.

Merlevede, B., M. de Zwaan, K. Lenaerts, and V. Purice (2015). Multinational networks, domestic and foreign firms in Europe. *Working Paper D/2015/7012/02, Working Papers of Faculty of Economics and Business Administration, Ghent University, Belgium.*

Merlevede, B. and B. Michel (2020). Downstream offshoring and firm-level employment. *Canadian Journal of Economics/Revue canadienne d'économique 53*(1), 249–283.

Michel, B. and F. Rycx (2014). Productivity gains and spillovers from offshoring. *Review of International Economics 22*(1), 73–85.

Moulton, B. R. (1986). Random group effects and the precision of regression estimates. *Journal of Econometrics 32*(3), 385–397.

Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *Review of Economics and Statistics 72*(2), 334–338.

OECD (Accessed on 06 June 2019). Gross domestic spending on R&D (indicator). URL: `https://data.oecd.org/rd/gross-domestic-spending-on-r-d.htm#indicator-chart`. DOI: `https://doi.org/10.1787/d8b068b4-en`.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6), 1263–1297.

Parente, S. L. (1994). Technology adoption, learning-by-doing, and economic growth. *Journal of Economic Theory 63*(2), 346–369.

Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics 71*(1), 135–158.

Shu, P. and C. Steinwender (2019). The impact of trade liberalization on firm productivity and innovation. *Innovation Policy and the Economy 19*, 39–68.

Solow, R. M. (1957). Technical change and the aggregate production function. *Review of Economics and Statistics 39*(3), 312–320.

Stokey, N. L. (1988). Learning by doing and the introduction of new goods. *Journal of Political Economy 96*(4), 701–717.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature 49*(2), 326–65.

Timmer, M. P., E. Dietzenbacher, B. Los, R. Stehrer, and G. J. de Vries (2015). An illustrated user guide to the World Input–Output Database: The case of global automotive production. *Review of International Economics 23*(3), 575–605.

Timmer, M. P., B. Los, R. Stehrer, and G. J. de Vries (2016). An anatomy of the global trade slowdown based on the WIOD 2016 release. Technical report, Groningen Growth and Development Centre, University of Groningen.

UNCTAD (Accessed on 05 June 2019). Country classification. URL: `https://unctadstat.unctad.org/EN/Classifications.html`.

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

Windmeijer, F. (2021). Testing underidentification in linear models, with applications to dynamic panel and asset pricing models. *Journal of Econometrics*.

A Appendix

Table A.1: List of CPA and NACE 2-digit (Rev.2) industries for the manufacturing sector

CPA	NACE	Description of CPA industries
5	10t12	Manufacture of food products, beverages and tobacco products
6	13t15	Manufacture of textiles, wearing apparel and leather products
7	16	Manufacture of wood and products of wood and cork, except furniture; etc.
8	17	Manufacture of paper and paper products
9	18	Printing and reproduction of recorder media
10	19	Manufacture of coke and refined petroleum products
11	20	Manufacture of chemicals and chemical products
12	21	Manufacture of basic pharmaceutical products and
13	22	Manufacture of rubber and plastic products
14	23	Manufacture of other non-metallic mineral products
15	24	Manufacture of basic metals
16	25	Manufacture of fabricated metal products, except machinery and equipment
17	26	Manufacture of computer, electronic and optical products
18	27	Manufacture of electrical equipment
19	28	Manufacture of machinery and equipment n.e.c.
20	29	Manufacture of motor vehicles, trailers and semi-trailers
21	30	Manufacture of other transport equipment
22	31t32	Manufacture of furniture; other manufacturing
23	33	Repair and installation of machinery and equipment

Note: CPA corresponds to industries in WIOD according to ISIC Rev. 4 or equivalently NACE Rev. 2 (Timmer et al. 2016).

Online Appendix*
Productivity Effects of Internationalisation through the Domestic Supply Chain

Bruno Merlevede† Angelos Theodorakopoulos‡

May 2021

*This document provides relevant supporting material for online publication.

†Department of Economics, Ghent University, Ghent, Belgium; e-mail: bruno.merlevede@ugent.be

‡Corresponding author: University of Oxford, Oxford Martin School, 34 Broad Street, Oxford OX1 3BD, United Kingdom; e-mail: angelos.theodorakopoulos@oxfordmartin.ox.ac.uk

A GNR Estimation Procedure

This section serves as an overview of the basic steps and assumptions needed to apply the non-parametric identification procedure of GNR when estimating equation (6) under a parametric specification of the production technology $\left(f_j(\cdot)\right)$ and firm fixed effects (ϕ_i). For a detailed and complete description refer to GNR. For simplicity and without loss of generality, we disregard the industry dimension j. The estimation discussed below is directly extended by allowing the functional form of the production technology $f(\cdot)$ to also vary by industry j. Therefore, under the presence of ϕ_i, the production function in (5) can be expressed as:[1]

$$y_{it} = f\left(k_{it}, l_{it}, m_{it}; \alpha\right) + \omega_{it} + \phi_i + \varepsilon_{it} \tag{A.1}$$

This case considers the classic environment of perfect competition in both input and output markets. Capital and labour are assumed to be predetermined inputs and therefore chosen one year prior to the realisation of productivity, ω_{it}, i.e. at $t-1$. The only flexible input in the specification is material, which is assumed to freely adjust in each period (variable) and have no dynamic implications (static).

Conditional on the state variables and other firm characteristics, a firm's static profit maximisation problem yields the first order condition with respect to the flexible input, material:

$$P_t^M = P_t \frac{\partial}{\partial M_t} F\left(K_{it}, L_{it}, M_{it}; \alpha\right) \mathrm{e}_{it}^{\omega} \mathscr{E} \tag{A.2}$$

where P_t^M and P_t is the price of material and output, respectively. Under perfect competition in input and output markets, they are constant across firms within the same country-industry but can vary across time. By the time firms make their annual decisions, ex-post shocks ε_{it} are not in their information set, and thus firms create expectations over them such that: $\mathscr{E} = E\left(\mathrm{e}^{\varepsilon_{it}}\right)$.

Combining the log of (A.2) with (A.1) and re-arranging terms, we retrieve a share equation:

$$s_{it} = ln\left(\widetilde{f}\left(k_{it}, l_{it}, m_{it}; \widetilde{\alpha}\right)\right) + ln\mathscr{E} - \varepsilon_{it} \tag{A.3}$$

where s_{it} is the log of the nominal share of material and $\widetilde{f}\left(k_{it}, l_{it}, m_{it}; \widetilde{\alpha}\right) = \frac{\partial}{\partial m_{it}} f\left(k_{it}, l_{it}, m_{it}; \alpha\right)$ is the output elasticity of the flexible input, material. Note that the share equation is net of the log additive TFP term ω_{it}, inducing the transmission bias, and the firm fixed effects ϕ_i.

A.1 Step One

A Non Linear Least Squares estimation of the share equation (A.3) is applied using the Gauss-Newton algorithm to minimise the sum of squared errors. Under a Cobb-Douglas production

[1] Given the data structure, we consider a large number of firms (N) and a small number of time series observations per firm (T). Thus, we rely on typical panel data asymptotic properties as $N \to \infty$ for fixed T.

technology $\widetilde{f}(k_{it}, l_{it}, m_{it}; \widetilde{\alpha})\mathscr{E} = \alpha_m \mathscr{E} \equiv \widetilde{\alpha}_m$, where α_m is now a constant representing the output elasticity of the flexible input material. This step identifies ε_{it} (hence $\mathscr{E} \equiv \sum_{it} \frac{\varepsilon_{it}}{NT}$) and $\widetilde{\alpha}_m$, which in turn allows us to compute $\alpha_m \equiv \widetilde{\alpha}_m / \mathscr{E}$.

A.2 Step Two

By integrating up the output elasticity of the flexible input:

$$\int \widetilde{f}(k_{it}, l_{it}, m_{it}; \widetilde{\alpha})\, dm_{it} = f(k_{it}, l_{it}, m_{it}; \alpha) + \mathscr{F}\left(k_{it}, l_{it}; \widetilde{\widetilde{\alpha}}\right) \tag{A.4}$$

we retrieve the production technology $f(\cdot)$ to the part that remains to be identified $\mathscr{F}\left(k_{it}, l_{it}; \widetilde{\widetilde{\alpha}}\right)$. By differencing it with the production function (A.1) we get the following expression for TFP:

$$\omega_{it} = \mathscr{Y}_{it} + \mathscr{F}\left(k_{it}, l_{it}; \widetilde{\widetilde{\alpha}}\right) - \phi_i \tag{A.5}$$

where $\mathscr{Y}_{it}$ is the log of the expected output net of the term (A.4) computed in Step One. Under a Cobb-Douglass production technology, $\mathscr{Y}_{it} = y_{it} - \widehat{\varepsilon}_{it} - \widehat{\alpha}_m m_{it}$ and $\mathscr{F}\left(k_{it}, l_{it}; \widetilde{\widetilde{\alpha}}\right) = \widetilde{\widetilde{\alpha}}_k k_{it} + \widetilde{\widetilde{\alpha}}_l l_{it}$, where $\widetilde{\widetilde{\alpha}}_k \equiv -\alpha_k$ and $\widetilde{\widetilde{\alpha}}_l \equiv -\alpha_l$.

To proceed, we combine the assumption over the law of motion of TFP used in baseline specification (6) with (A.5) to generate the following estimating equation:

$$\begin{aligned} \mathscr{Y}_{it} &= -\mathscr{F}\left(k_{it}, l_{it}; \widetilde{\widetilde{\alpha}}\right) + \rho_\omega \omega_{it-1} + \rho_p proxies_{jct-1} + \rho_x X_{jct-1} + \rho_{fe} d_{fe,t} + (1-\rho_\omega)\phi_i + \xi_{it} \\ &= -\mathscr{F}\left(k_{it}, l_{it}; \widetilde{\widetilde{\alpha}}\right) + \rho_\omega \left(\mathscr{Y}_{it-1} + \mathscr{F}\left(k_{it-1}, l_{it-1}; \widetilde{\widetilde{\alpha}}\right)\right) \\ &\qquad + \rho_p proxies_{jct-1} + \rho_x X_{jct-1} + \rho_{fe} d_{fe,t} + (1-\rho_\omega)\phi_i + \xi_{it} \end{aligned} \tag{A.6}$$

where $d_{fe,t}$ is a full set of dummies with their corresponding parameters ρ_{fe} representing the relevant time-varying fixed effects $\phi_{fe,t}$ in (6). In the absence of ϕ_i, one can readily estimate (A.6) using a Generalised Method of Moments (GMM) estimator (see GNR for more details). However, in the presence of firm fixed effects, further transformations and assumptions are necessary. We turn to this next.

A.2.1 First Difference GMM (DIF)

Following the dynamic panel literature, GNR augment their baseline estimator to account for firm fixed effects by first-differencing (A.6) such that:

$$\Delta\mathscr{Y}_{it} = -\Delta\mathscr{F}\left(k_{it}, l_{it}; \widetilde{\widetilde{\alpha}}\right) + \rho_\omega \Delta\omega_{it-1} + \rho_p \Delta proxies_{jct-1} + \rho_x \Delta X_{jct-1} + \rho_{fe} \Delta d_{fe,t} + \Delta\xi_{it} \tag{A.7}$$

where Δ is the first difference operator.[2] However, the above equation suffers from endogeneity induced by the correlation between $\Delta\omega_{it-1}$ and $\Delta\xi_{it}$. To solve for this, one could instrument with deeper lags in levels *à la* Arellano and Bond (1991). However, as discussed in section 3.3, such estimators are known to perform poorly due to weak instruments. Therefore, in the next section, we further augment the GNR estimator in the presence of firm fixed effects.

A.2.2 System GMM (SYS)

Following Blundell and Bond (1998), the SYS approach augments the DIF from the previous section by simultaneously estimating the equation in differences and levels:[3]

$$\begin{aligned}\begin{pmatrix}\Delta\mathscr{Y}_{it}\\ \mathscr{Y}_{it}\end{pmatrix} = & -\begin{pmatrix}\Delta\mathscr{F}\left(k_{it},l_{it};\tilde{\tilde{\alpha}}\right)\\ \mathscr{F}\left(k_{it},l_{it};\tilde{\tilde{\alpha}}\right)\end{pmatrix} + \rho_\omega\begin{pmatrix}\Delta\omega_{it-1}\\ \omega_{it-1}\end{pmatrix} + \rho_p\begin{pmatrix}\Delta proxies_{cjt-1}\\ proxies_{cjt-1}\end{pmatrix}\\ & + \rho_x\begin{pmatrix}\Delta X_{it-1}\\ X_{it-1}\end{pmatrix} + \rho_{fe}\begin{pmatrix}\Delta d_{fe,t}\\ d_{fe,t}\end{pmatrix} + \begin{pmatrix}\Delta\xi_{it}\\ \xi_{it}\end{pmatrix}\end{aligned} \tag{A.8}$$

where the same linear relationship with the same coefficients applies. This results in a stacked dataset with twice the number of firms and the same set of parameters used in levels.[4] By distinctly instrumenting each of the stacked equations, we form the $L \equiv \left(L^D + L^L\right)$ x 1 vector of stacked moment conditions:

$$E\left[m_i(\theta_o)\right] = E\left[\mathscr{Z}_i'\tilde{\xi}_i\right] = E\left[\begin{pmatrix}\mathscr{Z}_i^D & 0\\ 0 & \mathscr{Z}_i^L\end{pmatrix}'\begin{pmatrix}\Delta\xi_i\\ \xi_i\end{pmatrix}\right] = 0 \tag{A.9}$$

as a function of the K x 1 vector of unknown parameters $\theta_o = \left(\tilde{\tilde{\alpha}},\rho_\omega,\rho_p,\rho_x,\rho_{fe}\right)$ with $L > K$, where $\Delta\xi_i = (\Delta\xi_{i2},\ldots,\Delta\xi_{i,T})'$, $\xi_i = (\xi_{i1},\ldots,\Delta\xi_{i,T})'$, $\mathscr{Z}_i^D$ is a $T-1$ x L^D instrument matrix used to distinctly instrument the equation in first-differences, and $\mathscr{Z}_i^L$ is a T x L^L instrument matrix used to distinctly instrument the equation in levels.

The choice of the instruments is based on the timing assumptions of the variables and thus how they correlate with the error term, i.e. predetermined, endogenous, or exogenous. On the one hand, capital and labour are assumed to be predetermined inputs chosen at time $t-1$ and are thus uncorrelated to any current or future innovations of productivity. On the other hand, the proxies and additional controls which are treated as endogenous—correlated with contemporary but not future productivity innovations—are instrumented with (deeper) lags. While we rely

[2]Standard to dynamic panel methods, linearity in the law of motion of TFP is a necessary condition for eliminating ϕ_i under first-differences.

[3]This approach requires additional stationarity restrictions on the initial conditions process (Arellano and Bover 1995).

[4]In the first-differenced equation, the industry ($\phi_j t$) and country ($\phi_c t$) specific linear time trends included in $\phi_{fe,t}$ from the levels equation now enter in $\Delta d_{fe,t}$ as a set of industry (ϕ_j) and country fixed effects (ϕ_c), respectively. In addition, the vector $\Delta d_{fe,t}$ is extended with zeros to annihilate any time-invariant terms such as the constant and, in the case of industry-j specific production technology $f_j(\cdot)$, industry-j dummies.

on the first lag only for the firm-level controls, we exploit all available lag information for the country-industry-year-level proxies to maintain maximal identifying variation. Therefore, similar to the persistence term, for the equation in first differences, $\mathscr{Z}_i^D$ contains (deeper lag) values in levels (*à la* Arellano and Bond 1991). For the equation in levels, $\mathscr{Z}_i^L$ contains (lag) values in first-differences (*à la* Blundell and Bond 1998; Arellano and Bover 1995). Note that we exclude redundant instruments by choosing the first available lag in $\mathscr{Z}_i^D$ and all available lags in $\mathscr{Z}_i^L$ (see Appendix C from Kiviet et al. 2017, for a complete description on redundant instruments).

We abstain from using additional lag lengths for the firm-level controls to avoid potential biases generated by instrument proliferation (Roodman 2009). In the same spirit, we further limit the instrument count by using a collapsed version of the instrument matrix, as suggested by Roodman (2009) among others. Kiviet et al. (2017) and Kiviet (2020) demonstrate how the combination of these two instrument reduction methods, i.e. removing long lags and collapsing, can improve estimation precision.

Finally, the time-varying fixed effects $d_{fe,t}$ are assumed to be exogenous and thus orthogonal to the productivity shocks. To extract redundant instruments, we consider $d_{fe,t}$ only in $\mathscr{Z}_i^L$. It is important to mention here that $d_{fe,t}$ includes a constant which is by default exogenous and identifies the global mean.[5] Specifically, under a Cobb-Douglas production technology:

$$\mathscr{Z}_i^D = \begin{pmatrix} k_{i1} & l_{i1} & 0 & 0 & 0 \\ k_{i2} & l_{i2} & \mathscr{Y}_{i1} & proxies_{cj1} & X_{i1} \\ k_{i3} & l_{i3} & \mathscr{Y}_{i2} & proxies_{cj2} & X_{i2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ k_{i,T-1} & l_{i,T-1} & \mathscr{Y}_{i,T-2} & proxies_{cj,T-2} & X_{i,T-2} \end{pmatrix}$$

and

$$\mathscr{Z}_i^L = \begin{pmatrix} \Delta k_{i1} & \Delta l_{i1} & 0 & 0 & \cdots & 0 & 0 & d_{fe,1} \\ \Delta k_{i2} & \Delta l_{i2} & \Delta\mathscr{Y}_{i1} & \Delta proxies_{cj1} & & 0 & \Delta X_{i1} & d_{fe,2} \\ \Delta k_{i3} & \Delta l_{i3} & \Delta\mathscr{Y}_{i2} & \Delta proxies_{cj2} & & 0 & \Delta X_{i2} & d_{fe,3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & O & \vdots & \vdots \\ \Delta k_{iT} & \Delta l_{iT} & \Delta\mathscr{Y}_{i,T-1} & \Delta proxies_{cj,T-1} & \cdots & \Delta proxies_{cj1} & \Delta X_{iT-1} & d_{fe,T} \end{pmatrix}$$

This step proceeds with a GMM estimation which uses the sample analog of the population moment conditions (A.9) to construct an estimator for θ (Hansen 1982). The GMM estimator $\hat{\theta}$

[5] In our case of industry specific production technology $f_j(\cdot)$, we replace the constant with a full set of industry-j dummies which are assumed to be exogenous.

minimises the quadratic form:

$$J(\theta) = \left(\frac{1}{N}\sum_{i=1}^{N} m_i(\theta)\right)' W \left(\frac{1}{N}\sum_{i=1}^{N} m_i(\theta)\right) \tag{A.10}$$

with respect to θ where W is an L x L positive semi-definite weighting matrix. Given that the GMM objective function is of quadratic form, we solve for the minimum using the Gauss-Newton non-linear algorithm which involves iteration to convergence for a given W (one-step). Note that for inference we rely on bootstrapping and that for the SYS moment conditions there is no simple one-step efficient W. Therefore, as a choice of a suboptimal weighting matrix to yield a consistent one-step GMM estimator, we follow Blundell and Bond (2000) in setting $W = \left(\frac{1}{N}\sum_{i=1}^{N} \mathscr{Z}_i' H_i \mathscr{Z}_i\right)^{-1}$. This contains a block diagonal matrix $H_i = diag\left(D_i D_i', I_{T-1}\right)$, where D_i is a $T-1$ x T matrix with -1s in the diagonal, 1s in the first upper sub-diagonal, and zeros elsewhere, and I_{T-1} is an identity matrix of size $T-1$.[6]

By minimising the sample analogue of the GMM criterion function, we retrieve estimates for the remaining parameters of the production technology $\left(\widehat{\widehat{\widehat{\alpha}}}\right)$ as well as the Markov process parameters and the time varying fixed effects $\left(\widehat{\rho}_\omega, \widehat{\rho}_p, \widehat{\rho}_x, \widehat{\rho}_{fe}\right)$. For a Cobb-Douglas production technology the estimated production function is:

$$y_{it} = \widehat{\alpha}_k k_{it} + \widehat{\alpha}_l l_{it} + \widehat{\alpha}_m m_{it} + \omega_{it} + \phi_i + \widetilde{\varepsilon}_{it} \tag{A.11}$$

Using the estimated parameters from this two-step procedure, i.e. $\widehat{\widehat{\alpha}}$ from step one and $\widehat{\widehat{\widehat{\alpha}}}$ from step two, we can now compute productivity $\widehat{\omega_{it} + \phi_i}$ and other relevant functionals, e.g. returns to scale $RTS = \widehat{\alpha}_k + \widehat{\alpha}_l + \widehat{\alpha}_m$.

On a technical matter, it is important to mention that, within the GMM algorithm we 'net out' the exogenous fixed effects $d_{fe,t}$ using a partitioned regression (Frisch and Waugh 1933; Lovell 1963; Giles 1984). This approach reduces the parameter space that the GMM algorithm needs to search over, which has two empirical advantages. First, it exponentially reduces estimation time given both the iterative nature of the GMM algorithm and the bootstrapping procedure used to obtain standard errors. Second, we find empirically that it helps to avoid possible non-convergence issues of the estimator related to the presence of local-minima and flat regions in the criterion function.

Implementation of the partitioning is straightforward. Within each iteration of the GMM algorithm, we use the moment conditions related to the fixed effects to retrieve Ordinary Least

[6] Alternatively, for the consistent estimation of this one-step GMM estimator, one could use the suggestion from Windmeijer (2000), where the lower-left and upper-right zero quadrants of matrix H_i are replaced by matrix D_i' and D_i, respectively. Or, more simply, one could follow Arellano and Bover (1995) and Blundell and Bond (1998) in setting $W = \left(\frac{1}{N}\sum_{i=1}^{N} \mathscr{Z}_i' \mathscr{Z}_i\right)^{-1}$. For a detailed overview on this topic, see Kiviet et al. (2017) and the Online Appendix of Kripfganz and Schwarz (2019).

Squares (OLS) estimates for ρ_{fe}. Specifically, in each iteration, for a given a set of starting values for the parameters θ other than ρ_{fe},[7] we regress the left hand side variable minus the right-hand side part of equation (A.8)—excluding the part related to the fixed effects—on the set of fixed effects. In turn, the ρ_{fe} 'OLS estimates' are used in (A.8) to complete the GMM iteration under the full set of instruments $\mathscr{Z}_i$. We repeat this approach within each iteration of the GMM algorithm until convergence is achieved. As such, we can now calculate the time-consuming cross-products and inversions of large matrices needed for the OLS outside of the iterative procedure. Also, the GMM parameter space is reduced drastically. For example, in our baseline model, from a total of 140 parameters we now need to estimate through the computationally intensive GMM only 49 parameters since the additional 91 parameters related to the fixed effects are partitioned and obtained from a computationally fast OLS regression.

[7] We use OLS estimates of equation (A.8).

B Over- and Underidentification Test Statistics

In this section we describe the construction of the relevant statistics to test the model for overidentification and underidentification. The construction of the bootstrap p-values for the test statistics is discussed next in the Online Appendix C.3.

Overidentification.—We test the validity of overidentifying moment conditions using the Hansen-J test (Hansen 1982). For the construction of the Hansen-J statistic we use $J = NJ(\widehat{\theta})$, which is computed using the consistent one-step GMM estimates $\widehat{\theta}$ and the weight matrix specified as before.

Underidentification.—We test for weak identification using the underidentification test established by Windmeijer (2021) for models with complex data structures, i.e. clustered and potentially heteroskedastic dynamic panel data models estimated by GMM. This test builds upon the simple test for weak instruments by Sanderson and Windmeijer (2016). Windmeijer (2021) highlights that the underidentification test is equivalent to an overidentification test when regressing any endogenous variable on the remaining regressors of the original model using the same set of instruments. For the case of the k^{th} proxy ($proxies^k_{cjt-1}$), the auxiliary model is:

$$\begin{aligned} \begin{pmatrix} \Delta proxies^k_{cjt-1} \\ proxies^k_{cjt-1} \end{pmatrix} = - \begin{pmatrix} \Delta \mathscr{F}\left(k_{it}, l_{it}; \widetilde{\widetilde{\alpha}}\right) \\ \mathscr{F}\left(k_{it}, l_{it}; \widetilde{\widetilde{\alpha}}\right) \end{pmatrix} + \rho_\omega \begin{pmatrix} \Delta\omega_{it-1} \\ \omega_{it-1} \end{pmatrix} + \rho_p^s \begin{pmatrix} \Delta proxies^{-k}_{cjt-1} \\ proxies^{-k}_{cjt-1} \end{pmatrix} \\ + \rho_x \begin{pmatrix} \Delta X_{it-1} \\ X_{it-1} \end{pmatrix} + \rho_{fe} \begin{pmatrix} \Delta d_{fe,t} \\ d_{fe,t} \end{pmatrix} + \begin{pmatrix} \Delta\xi_{it} \\ \xi_{it} \end{pmatrix} \end{aligned} \tag{B.1}$$

where the dependent variable from the original model is replaced with $proxies^k_{cjt-1}$. In turn, this is excluded from the set of regressors ($proxies^{-k}_{cjt-1}$). The same estimation as the original model is thus performed while keeping the instrument matrix $\mathscr{Z}_i$ unchanged. Overall, for each endogenous explanatory variable, the overidentification Hansen-J test for the relevant auxiliary model serves as an underidentification test under the null hypothesis that the model is underidentified. The test relies on the choice of the left-hand side variable, and thus can only inform whether the particular endogenous variable is poorly predicted by the instruments. The Hansen-J test statistic is computed in the same way as above.

C Bootstrap

In this section we discuss how to operationalise the cluster bootstrap procedure for our dynamic panel data model estimated from an overidentified GMM. In turn, we show how to compute cluster bootstrap standard errors and p-values for two model specification tests, i.e. overidentification and underidentification tests.

C.1 Implementation and Calculation of Standard Errors

We implement a cluster bootstrap where we first define clusters G at the industry-country level, i.e. firm-year observations can be arbitrarily correlated within but independent across clusters.[8] We form clusters at the industry-country and not at the industry-country-year level of our regressors of interest in order to ensure that—given the dynamic representation of our model—the full time-series of each firm is retained when creating the bootstrap samples below (Horowitz 2001). If anything, this choice allows for a more flexible error structure, whereby errors within the cluster can also be arbitrarily correlated over time. We then randomly draw with replacement G times over entire clusters, i.e. blocks of firms, from the original sample to generate the b^{th} bootstrap sample, where $b = 1 \dots B$.[9] We repeat this exercise for $B = 99$ times and for each parameter estimate from the original sample $\widehat{\theta}$, $\widehat{\theta}_b$ is the estimate from the b^{th} bootstrap replication and $\bar{\theta}$ is the mean of all the $\widehat{\theta}_b$s. As such, the bootstrap standard error can be written as:

$$se(\widehat{\theta}) = \left(\frac{1}{B-1} \sum_{b=1}^{B} \left(\widehat{\theta}_b - \bar{\theta} \right)^2 \right)^{1/2} \tag{C.1}$$

Calculated as such, the computed standard errors can be used for statistical inference similar to any other asymptotically valid standard errors.

C.2 Recentering

For reliable bootstrap inference and testing of the over-identified GMM estimation described in Online Appendix A.2.2, we follow Hall and Horowitz (1996) to recenter the bootstrap moment conditions.[10] Specifically, for each b^{th} bootstrap sample, the GMM estimator $\widehat{\theta}_b$ minimises the following criterion function:

$$\widetilde{J}(\theta) = \left(\frac{1}{N} \sum_{i^b=1}^{N} \left(m_{i^b}(\theta) - \frac{1}{N} \sum_{i=1}^{N} m_i(\widehat{\theta}) \right) \right)' W_b \left(\frac{1}{N} \sum_{i^b=1}^{N} \left(m_{i^b}(\theta) - \frac{1}{N} \sum_{i=1}^{N} m_i(\widehat{\theta}) \right) \right) \tag{C.2}$$

[8] We form a total of 360 clusters for this application. This comes from the fact that we have 19 industries in each of the 19 countries, and exclude CPA classification 10 in Norway (due to a lack of data).

[9] Note that we use a pairs bootstrap, i.e. draw pairs of Y (left-hand side variable), X (right-hand side variables).

[10] See Bond and Windmeijer (2005) for such an application when comparing the finite sample performance of various test procedures for a range of dynamic panel data models using GMM. Alternatively, one could follow Brown and Newey (2002) by drawing bootstrap samples under a specific weighting of the original data ensuring that the moment conditions hold.

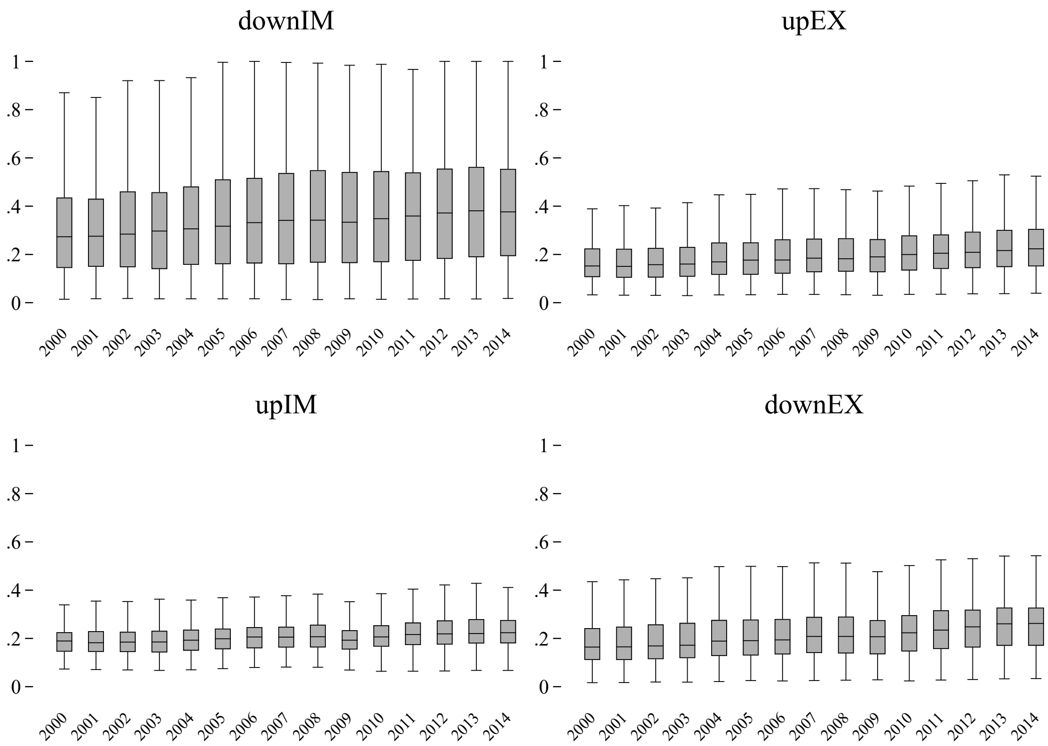
where the bootstrap moment conditions are recentered relative to the original sample moment conditions under the consistent one-step GMM estimates from the original sample $\widehat{\theta}$. W_b is constructed similarly to the weighting matrix used in (A.10) under the bootstrap sample. As such, the bootstrap version of the Hansen-J statistic is now based on $J_b = N\widetilde{J}(\widehat{\theta}_b)$.

C.3 p-values for Over- and Underidentification Tests

For each bootstrap sample b we calculate the relevant test statistic J_b and create its bootstrap empirical distribution. Recall that the underidentification test is equivalent to the overidentification test where J_b is the bootstrap Hansen-J statistic from the auxiliary model outlined in Online Appendix B. The percentile in the bootstrap distribution of J_b is then given by $p_J = \frac{1}{B}\sum_{b=1}^{B} \mathbb{1}(J_b > J)$, where the indicator function is equal to one each time the bootstrap sample statistic J_b is strictly larger than the original sample statistic J, and 0 otherwise. If $p_J < \alpha$, the test rejects the null hypothesis—at size α.

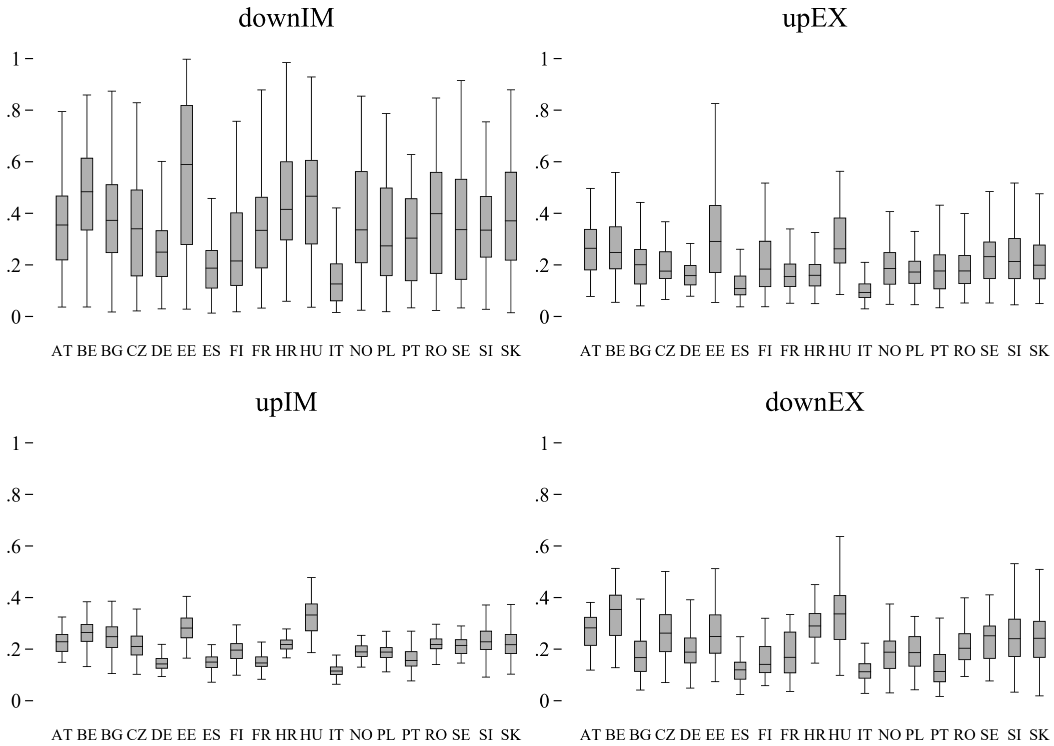
D Additional Figures and Tables

Figure D.1: Inter-industry importing and exporting by year



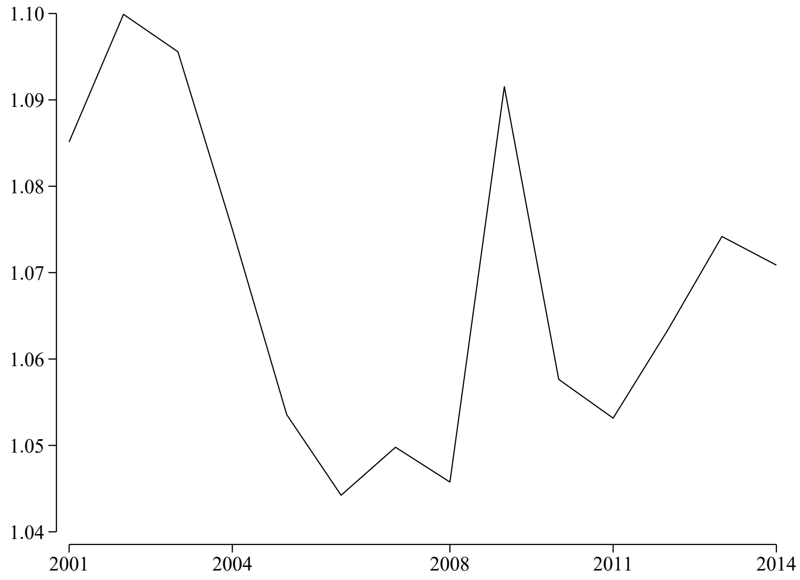
Source: Authors' calculations based on WIOT.
Notes: Let x represent the variable of interest. The upper, middle and lower hinge of the box represents the 25th ($x_{[25]}$), 50th ($x_{[50]}$) and 75th ($x_{[75]}$) percentile, respectively. Define $x_{(i)}$ as the ith ordered value of x. The upper adjacent line has a value $x_{(i)}$ such that $x_{(i)} \leq U$ and $x_{(i+1)} > U$, where $U = x_{[75]} + 1.5(x_{[75]} - x_{[25]})$. The lower adjacent line has a value $x_{(i)}$ such that $x_{(i)} \geq L$ and $x_{(i+1)} < L$, where $L = x_{[25]} - 1.5(x_{[75]} - x_{[25]})$.

Figure D.2: Inter-industry importing and exporting by country



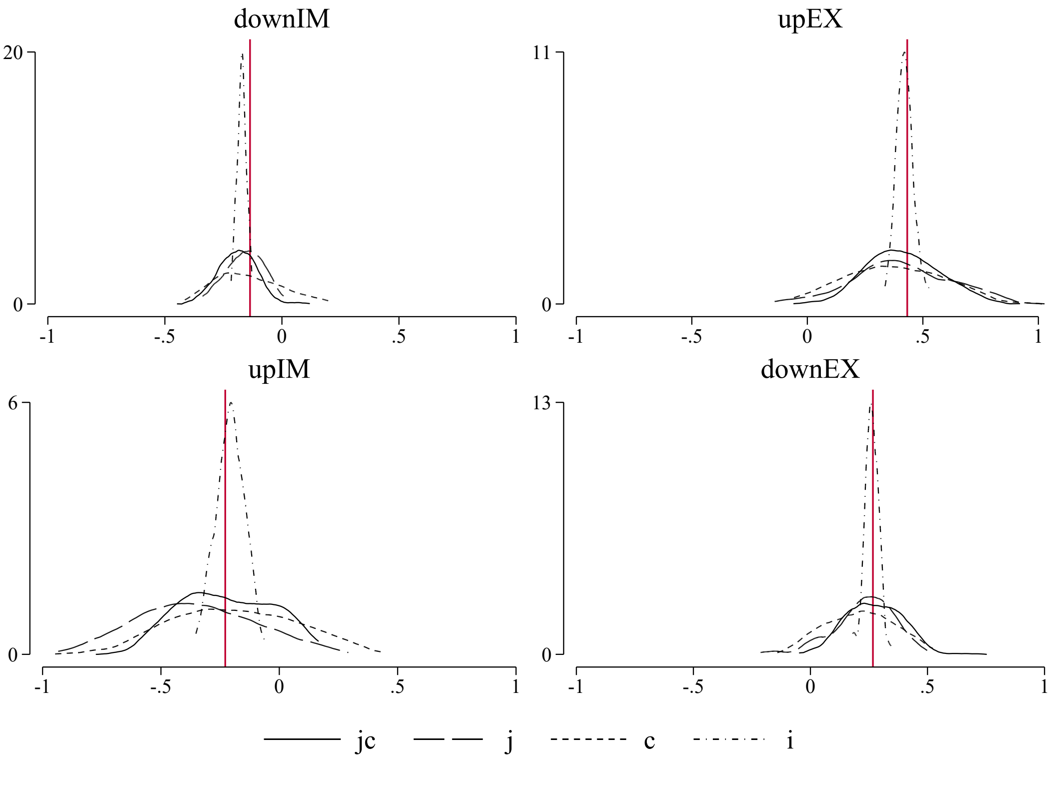
Source: Authors' calculations based on WIOT.
Notes: Let x represent the variable of interest. The upper, middle and lower hinge of the box represents the 25th ($x_{[25]}$), 50th ($x_{[50]}$) and 75th ($x_{[75]}$) percentile, respectively. Define $x_{(i)}$ as the ith ordered value of x. The upper adjacent line has a value $x_{(i)}$ such that $x_{(i)} \leq U$ and $x_{(i+1)} > U$, where $U = x_{[75]} + 1.5(x_{[75]} - x_{[25]})$. The lower adjacent line has a value $x_{(i)}$ such that $x_{(i)} \geq L$ and $x_{(i+1)} < L$, where $L = x_{[25]} - 1.5(x_{[75]} - x_{[25]})$.

Figure D.3: Markup by year



Source: Authors' calculations based BvDEP data and the estimated extension of the baseline model (6) assuming monopolistic competition in the output market and CES preferences.

Figure D.4: Distributions of bootstraped values for different clustering levels



Source: Authors' calculations based on BvDEP.
Notes: For each proxy variable of interest, the plotted distributions represent the kernel densities of the point estimates from the 99 replications of the cluster bootstrap for different types of clustering, i.e. industry-country (jc), industry (j), country (c), and firm (i). The vertical (red) line represents the point estimate of each variable from the original sample.

Table D.1: Summary Statistics for sample selection criteria

Sample Selection Criteria	(1) Observations	(2) # Firms	(3) Sales	(4) # Employees
1. Active Legal Status	99.74	99.98	99.94	99.86
2. Consolidated Accounts	95.82	96.48	88.23	82.47
3. No Missing Data	61.85	71.82	70.46	75.14
4. >20 Employees	30.66	23.70	92.43	89.29
5.i. BACON 30th percentile	93.84	94.85	78.44	88.71
5.ii. >2 consecutive observations	94.24	76.97	93.29	95.37
6.i. BACON 15th percentile	97.82	98.25	94.37	97.12
6.ii. >2 consecutive observations	94.42	77.48	89.33	95.14

Notes: This table reports the remaining percentage coverage of the firm-level sample across four different categories (columns 1-4), after applying each sample selection criterion (in each row). Each selection criterion is applied sequentially and thus the table reads from top to bottom rows. For the first two criteria, results are reported relative to the original sample. For criterion 5.i and 6.i, results are relative to the sample after the first four criteria are applied, and for criteria 5.ii and 6.ii results are relative to the sample after the application up to criterion 5.i and 6.i, respectively.

Table D.2: Output elasticities of inputs and returns to scale

CPA	Baseline				Imperfect Competition			
Industry	α_k	α_l	α_m	RTS	α_k	α_l	α_m	RTS
5	0.162	0.119	0.519	0.800	0.133	0.108	0.555	0.796
6	0.199	0.229	0.339	0.767	0.206	0.241	0.362	0.809
7	0.192	0.174	0.487	0.853	0.167	0.171	0.521	0.858
8	0.134	0.142	0.478	0.753	0.125	0.141	0.510	0.776
9	0.141	0.204	0.325	0.671	0.150	0.218	0.347	0.714
10	0.189	0.127	0.537	0.853	0.076	0.053	0.573	0.702
11	0.152	0.130	0.476	0.758	0.154	0.137	0.508	0.799
12	0.113	0.189	0.351	0.653	0.118	0.205	0.375	0.698
13	0.184	0.187	0.461	0.831	0.182	0.191	0.492	0.865
14	0.188	0.250	0.390	0.828	0.194	0.265	0.417	0.875
15	0.166	0.190	0.476	0.832	0.174	0.199	0.508	0.881
16	0.177	0.246	0.354	0.776	0.214	0.275	0.378	0.867
17	0.146	0.215	0.385	0.746	0.182	0.245	0.411	0.838
18	0.173	0.159	0.445	0.777	0.195	0.178	0.475	0.849
19	0.164	0.212	0.410	0.786	0.202	0.244	0.438	0.884
20	0.168	0.218	0.493	0.879	0.194	0.242	0.526	0.962
21	0.201	0.244	0.373	0.818	0.219	0.258	0.398	0.875
22	0.163	0.228	0.415	0.806	0.157	0.235	0.444	0.835
23	0.203	0.218	0.293	0.714	0.227	0.250	0.313	0.790
All-mean	0.172	0.200	0.416	0.788	0.182	0.213	0.445	0.839
All-median	0.173	0.212	0.410	0.786	0.194	0.241	0.438	0.849
All-st.dev.	0.018	0.044	0.066	0.041	0.030	0.056	0.070	0.048

Notes: α_k, α_l, α_m are point estimates of the output elasticities of capital, labour and material, respectively, under the baseline model (Baseline) and an extension accounting for monopolistic competition in output and CES preferences (Imperfect Competition). $RTS = \alpha_k + \alpha_l + \alpha_m$ is the returns to scale of production. The last three rows report the mean, median and standard deviation of the point estimates across all industries.

Table D.3: TFP effects from inter-industry importing and exporting under baseline specification with alternative combinations of proxies

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$downIM_{jct-1}$	-0.136	-0.083	-0.001				0.052			
	(0.091)	(0.066)	(0.111)				(0.068)			
$upEX_{jct-1}$	0.432***	0.370**		0.389**				0.360**		
	(0.166)	(0.154)		(0.175)				(0.149)		
$upIM_{jct-1}$	-0.228	-0.086			0.205				0.348	
	(0.207)	(0.224)			(0.208)				(0.221)	
$downEX_{jct-1}$	0.267**	0.304**				0.304*				0.337**
	(0.134)	(0.126)				(0.174)				(0.142)
IM_{jct-1}	0.254**		0.408***	0.279*	0.346**	0.356**				
	(0.117)		(0.146)	(0.158)	(0.164)	(0.172)				
EX_{jct-1}	0.002		-0.069	-0.078	-0.057	-0.102				
	(0.142)		(0.177)	(0.154)	(0.145)	(0.162)				
Observations	1,018,643	1,018,643	1,018,643	1,018,643	1,018,643	1,018,643	1,018,643	1,018,643	1,018,643	1,018,643

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. All regressions include: the persistence term; dummies for domestic and foreign ownership links; year, industry and country fixed effects; and industry and country linear time trends. Standard errors are computed using a cluster (at the industry-country) bootstrap with 99 replications and are reported in parentheses below point estimates.

Table D.4: TFP effects from inter-industry importing and exporting under baseline specification and robustness to alternative assumptions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
		Proxies			Size		BACON	Bootstrap		Clustering	
	Baseline	0 diagonal	Broad	Small	Large	Dummy	Trim p15	499 Reps	j-level	c-level	i-level
$downIM_{jct-1}$	-0.136	-0.123	-0.283	-0.122	-0.040	-0.112	-0.135	-0.136	-0.136*	-0.136	-0.136***
	(0.091)	(0.087)	(0.256)	(0.110)	(0.091)	(0.088)	(0.119)	(0.097)	(0.082)	(0.143)	(0.020)
$upEX_{jct-1}$	0.432***	0.415***	0.780***	0.710***	0.273*	0.492***	0.527**	0.432***	0.432*	0.432**	0.432***
	(0.166)	(0.144)	(0.262)	(0.244)	(0.148)	(0.166)	(0.226)	(0.163)	(0.224)	(0.220)	(0.039)
$upIM_{jct-1}$	-0.228	0.072	-0.403	-0.376	-0.364*	-0.251	-0.106	-0.228	-0.228	-0.228	-0.228***
	(0.207)	(0.295)	(0.288)	(0.311)	(0.207)	(0.198)	(0.285)	(0.231)	(0.268)	(0.291)	(0.061)
$downEX_{jct-1}$	0.267**	0.114	0.341*	0.153	0.168	0.255**	0.329*	0.267**	0.267*	0.267*	0.267***
	(0.134)	(0.102)	(0.186)	(0.143)	(0.110)	(0.127)	(0.175)	(0.128)	(0.138)	(0.149)	(0.030)
$D * downIM_{jct-1}$						0.152*					
						(0.092)					
$D * upEX_{jct-1}$						-0.289					
						(0.216)					
$D * upIM_{jct-1}$						-0.349					
						(0.255)					
$D * downEX_{jct-1}$						-0.121					
						(0.175)					
Observations	1,018,643	1,018,643	1,018,643	435,223	583,420	1,018,643	1,065,512	1,018,643	1,018,643	1,018,643	1,018,643

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. All regressions include: the persistence term; intra-industry importing and exporting; dummies for domestic and foreign ownership links; year, industry and country fixed effects; and industry and country linear time trends. D is a dummy variable equal to one when firms are classified as small-sized and zero when medium-large-sized. D is not reported since it is time invariant and thus not separately identified from the fixed effects. Column 4 regression also includes the interactions of intra-industry importing and exporting with D. Standard errors are computed using a cluster (at the industry-country—except for columns 7-9 at the industry, country, and firm, respectively) bootstrap with 99 replications (except for column 5 with 499 replications) and are reported in parentheses below point estimates.

Table D.5: Heterogeneity in TFP effects from inter-industry importing and exporting with sample splits

	(1) Baseline	(2) Domestic Links	(3) Foreign Links	(4) Relatively Downstream	(5) Relatively Upstream	(6) Relatively Low-tech	(7) Relatively High-tech	(8) Labour Intensive	(9) Capital Intensive	(10) Downstream & Low-tech	(11) Downstream & High-tech	(12) Downstream & Labour-int	(13) Downstream & Capital-int
$downIM_{jct-1}$	-0.136	-0.171	-0.013	-0.214**	-0.146	-0.140	-0.155	-0.333**	-0.229*	-0.120	-0.230*	-0.126	-0.248**
	(0.091)	(0.107)	(0.063)	(0.095)	(0.166)	(0.128)	(0.110)	(0.151)	(0.121)	(0.154)	(0.118)	(0.141)	(0.118)
$upEX_{jct-1}$	0.432***	0.556***	0.036	0.680***	-0.163	1.176***	0.044	1.031***	0.163	1.427***	0.143	1.421***	0.182
	(0.166)	(0.183)	(0.081)	(0.184)	(0.295)	(0.266)	(0.163)	(0.342)	(0.140)	(0.399)	(0.176)	(0.315)	(0.203)
$upIM_{jct-1}$	-0.228	-0.302	-0.112	-0.408	0.379	-1.144***	-0.062	-0.698*	-0.138	-1.309*	-0.390	-1.336**	-0.441
	(0.207)	(0.241)	(0.126)	(0.396)	(0.272)	(0.439)	(0.249)	(0.360)	(0.292)	(0.711)	(0.459)	(0.566)	(0.480)
$downEX_{jct-1}$	0.267**	0.236*	0.262***	0.188	0.346*	0.240	0.333**	0.281	0.177	0.268	0.149	0.217	0.082
	(0.134)	(0.123)	(0.073)	(0.155)	(0.194)	(0.173)	(0.143)	(0.185)	(0.167)	(0.205)	(0.183)	(0.224)	(0.176)
Observations	1,018,643	890,661	84,461	533,966	484,676	403,163	615,479	608,339	410,304	305,033	228,933	317,264	216,702

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. All regressions include: the persistence term; intra-industry importing and exporting; dummies for domestic and foreign ownership links; year, industry and country fixed effects; and industry and country linear time trends. Column 2 and 3 do not include SHH^{for}, SUB^{for} and SHH^{dom}, SUB^{dom}, respectively, as these variables are used as criteria to split the sample. Standard errors are computed using a cluster (at the industry-country) bootstrap with 99 replications and are reported in parentheses below point estimates.

Table D.6: TFP effects from inter-industry importing and exporting under baseline specification and robustness to alternative assumptions

	(1)	(2)	(3)
		D=1 if Low-tech	
	Baseline	CEEC	WEC
$downIM_{jct-1}$	-0.136	0.134	-0.083
	(0.091)	(0.161)	(0.099)
$upEX_{jct-1}$	0.432***	0.124	-0.265
	(0.166)	(0.200)	(0.172)
$upIM_{jct-1}$	-0.228	0.043	1.021***
	(0.207)	(0.275)	(0.332)
$downEX_{jct-1}$	0.267**	-0.012	0.393***
	(0.134)	(0.164)	(0.134)
$D * downIM_{jct-1}$		-0.346	-0.249
		(0.227)	(0.160)
$D * upEX_{jct-1}$		0.875**	0.610
		(0.386)	(0.437)
$D * upIM_{jct-1}$		-0.933	-0.024
		(0.587)	(0.486)
$D * downEX_{jct-1}$		0.511	-0.238
		(0.361)	(0.243)
Observations	1,018,643	277,003	741,640

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. All regressions include: the persistence term; intra-industry importing and exporting; dummies for domestic and foreign ownership links; year, industry and country fixed effects; and industry and country linear time trends. D is a dummy variable equal to one when firms are in Low-tech industries and zero otherwise. D is not reported since it is time invariant and thus not separately identified from the fixed effects. Column 2 regression also includes the interactions of intra-industry importing and exporting with D. CEEC refers to Cenrtal Eastern European Countries while WEC refers to Western European Countries. Standard errors are computed using a cluster (at the industry-country) bootstrap with 99 replications and are reported in parentheses below point estimates.

Table D.7: Upstreamness measure

Production Line Position	CPA	(1) *Mean EU*	(2) *Mean*
1	12	1.42	1.29
2	22	1.47	1.35
3	21	1.48	1.38
4	17	1.52	1.24
5	20	1.55	1.21
6	6	1.56	1.29
7	5	1.59	1.48
8	19	1.60	1.31
9	18	1.87	1.34
10	10	1.98	2.42
11	23	2.03	2.02
12	11	2.17	1.50
13	13	2.31	1.70
14	16	2.32	1.86
15	7	2.38	1.87
16	14	2.38	1.98
17	8	2.45	1.70
18	15	2.64	1.60
19	9	2.84	2.55

Notes: The upstreamness measures are computed as in Fally (2012) and Antràs et al. (2012) using WIOT. In column 1, we consider EU as one economy and, thus, for each industry-year we use the sum of WIOT tables across all 19 EU countries to construct the measures. In column 2, we use all available granular information to compute the measures, i.e. each industry-country-year WIOT tables separately. *Mean EU* is the per industry mean of the computed EU wide upstreamness measure across time. *Mean* is the per industry mean of the upstreamness measure across all EU countries and time. Larger values represent more upstream industries.

References

Antràs, P., D. Chor, T. Fally, and R. Hillberry (2012). Measuring the upstreamness of production and trade flows. *American Economic Review: Papers & Proceedings 102*(3), 412–416.

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies 58*(2), 277–297.

Arellano, M. and O. Bover (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics 68*(1), 29–51.

Blundell, R. and S. Bond (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics 87*(1), 115–143.

Blundell, R. and S. Bond (2000). GMM estimation with persistent panel data: an application to production functions. *Econometric Reviews 19*(3), 321–340.

Bond, S. and F. Windmeijer (2005). Reliable inference for GMM estimators? Finite sample properties of alternative test procedures in linear panel data models. *Econometric Reviews 24*(1), 1–37.

Brown, B. W. and W. K. Newey (2002). Generalized Method of Moments, efficient bootstrapping, and improved inference. *Journal of Business & Economic Statistics 20*(4), 507–517.

Fally, T. (2012). Production staging: Measurement and facts. *Boulder, Colorado, University of Colorado Boulder*, 155–168.

Frisch, R. and F. V. Waugh (1933). Partial time regressions as compared with individual trends. *Econometrica 1*(4), 387–401.

Giles, D. E. A. (1984). Instrumental variables regressions involving seasonal data. *Economics Letters 14*(4), 339–343.

Hall, P. and J. L. Horowitz (1996). Bootstrap critical values for tests based on generalized-method-of-moments estimators. *Econometrica 64*(4), 891–916.

Hansen, L. P. (1982). Large sample properties of Generalized Method of Moments estimators. *Econometrica 50*(4), 1029–1054.

Horowitz, J. L. (2001). Chapter 52 - The bootstrap. Volume 5 of *Handbook of Econometrics*, pp. 3159–3228. Elsevier.

Kiviet, J., M. Pleus, and R. Poldermans (2017). Accuracy and efficiency of various GMM inference techniques in dynamic micro panel data models. *Econometrics 5*(1).

Kiviet, J. F. (2020). Microeconometric dynamic panel data methods: Model specification and selection issues. *Econometrics and Statistics 13*, 16 – 45.

Kripfganz, S. and C. Schwarz (2019). Estimation of linear dynamic panel data models with time-invariant regressors. *Journal of Applied Econometrics*, 1–21.

Lovell, M. C. (1963). Seasonal adjustment of economic time series and multiple regression analysis. *Journal of the American Statistical Association 58*(304), 993–1010.

Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics 71*(1), 135–158.

Sanderson, E. and F. Windmeijer (2016). A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics 190*(2), 212–221. Endogeneity Problems in Econometrics.

Windmeijer, F. (2000). *Efficiency Comparisons for a System GMM Estimator in Dynamic Panel Data Models*, pp. 175–184. Boston, MA: Springer US.

Windmeijer, F. (2021). Testing underidentification in linear models, with applications to dynamic panel and asset pricing models. *Journal of Econometrics*.