

Domestic value chains as stepping stones to global value chain integration*

Cosimo Beverelli[†]

Robert B. Koopman^{†‡}

Victor Stolzenburg^{†‡}

Simon Neumueller[§]

January 12, 2019

Abstract

We study the role of Domestic Value Chains (DVCs) for Global Value Chain (GVC) integration. In the presence of industry specific fixed costs of fragmenting production and of switching across input suppliers, DVCs can either be stepping stones or stumbling blocks for GVCs. Focusing on backward linkages, that is the sourcing of intermediates, we provide robust empirical evidence in favour of the stepping stone hypothesis. In our benchmark specification a one standard deviation increase in DVC integration raises subsequent GVC integration by about 0.4%. To identify the mechanisms at work, we exploit two dimensions of industry level heterogeneity: product differentiation (a proxy of fragmentation costs) and relationship specificity (a proxy of the costs of switching between suppliers). We find that DVC integration is less conducive to GVC integration in industries that are characterized by relatively high switching costs and relatively low fragmentation costs.

Keywords: Domestic value chains; global value chains; input-output linkages.

JEL Classification: F14; F15; F63.

*The opinions expressed in this paper should be attributed to its authors. They are not meant to represent the positions or opinions of the WTO and its Members and are without prejudice to Members' rights and obligations under the WTO. Any errors are attributable to the authors. This paper builds upon background research for the Global Value Chain Development Report 2017, a joint project by IDE-JETRO, OECD, RCGVC, World Bank, and WTO. Without implicating them, we thank seminar participants at the GVC Development Report 2016 Background Paper Conference, the Geneva Trade and Development Workshop, the University of Trento, the 16th RIEF meetings, and the Graduate Institute, Geneva for useful comments.

[†]Economic Research Division, World Trade Organization. Rue de Lausanne 154, 1202 Geneva, Switzerland. E-mail: cosimo.beverelli@wto.org; robert.koopman@wto.org; victor.stolzenburg@wto.org (corresponding author).

[‡]Graduate Institute of International and Development Studies, Chemin Eugène-Rigot 2, 1202 Geneva, Switzerland.

[§]Office for West Africa, United Nations Economic Commission for Africa, P.O. BOX 744 Niamey, Niger. E-mail: NSimon@uneca.org.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/twec.12779

This article is protected by copyright. All rights reserved.

1 Introduction

Global value chains (GVCs) are an important phenomenon of 21st century trade. Not only final goods or services get traded, but intermediate products and stages get outsourced and production becomes more and more fragmented. GVCs are often developed by large firms that coordinate input sourcing and assembly decisions, establishing industry linkages across borders.¹ Seminal work by Hummels et al. (2001) unveiled the growing importance of this international production sharing. In more recent work, Johnson and Noguera (2012a, 2012b) characterize the difference between value added trade and gross trade, showing that the GVC revolution, as measured by trade in value added, is ongoing. In a similar fashion, Timmer et al. (2014) show that global fragmentation, proxied by the foreign value added content of production, has rapidly increased since the early 1990s.

This development is of considerable interest for policy makers since both theory and empirics suggest that integrating in GVCs can lead to higher productivity and GDP. For instance, Baldwin and Robert-Nicoud (2014) show how GVCs cause productivity improvements akin to technological change by embedding the prominent features of the Grossman and Rossi-Hansberg (2008) model into a general equilibrium setting. Empirically, Kummritz (2016) finds that at the industry level higher GVC integration results in higher labor productivity and value added.

A central question that arises in this context is what determines GVC integration. Hummels et al. (2001), Johnson and Noguera (2012a), and Baldwin and Lopez Gonzalez (2015) find that structural factors such as country size, industrial structure, and location play a key role. In particular, small economies that are located close to GVC hubs such as Germany, Japan, and the United States exhibit strong linkages into GVCs. Kowalski et al. (2015) show that in addition policy matters. Countries with open trade and investment policies and sound institutions tend

¹ In the remainder, a distinction will be drawn between ‘industries’, ‘broad industries’ and ‘sectors’. Industries will refer to the ISIC Rev. 3 two digit codes (or aggregation thereof) listed in the first column of Table A-2. Broad industries will refer to the following six aggregates of industries, listed in the last column of Table A-2: Primary, Raw material processing, Light manufacturing, Heavy manufacturing, Electronics and Services. Sectors will be the three macro aggregates of economic activity: primary (ISIC codes 01-14), manufacturing (ISIC codes 15-37), and services (ISIC codes 40-95).

to integrate into GVCs more easily.

A substantial part of the variation in GVC integration, however, remains unexplained. To provide an illustration, Figure 1 plots the levels of GVC integration (measured as foreign value added in exports as a share of total exports) of three country pairs: Sweden and Belgium; Switzerland and Singapore; Romania and Bulgaria. Within each pair, countries are similar in terms of the structural and policy factors used in the literature to explain GVC participation. One can still see significant differences in levels of GVC integration within pairs of matched countries, especially in the last two pairs. Such differences cannot be explained by the drivers of GVC integration identified in the literature.

< Figure 1 about here >

A recent and growing strand of research, revisiting earlier work by Hirschman (1958), discusses the positive implications of domestic linkages within and across industries.² Building on this literature, in this paper we investigate whether such domestic linkages are stepping stones or stumbling blocks for GVCs, and the mechanisms behind the relationship between GVCs and domestic value chains (DVCs). Various papers have modeled firms' sourcing choices as either one between global offshoring and in-house production,³ or a simultaneous decision between global offshoring, domestic outsourcing and in-house production.⁴ Our interest is to empirically assess a sequential development, in which the amount of globally sourced intermediates is dependent on the pre-existing level of domestic outsourcing because fragmenting further and replacing domestic for foreign suppliers may be costly. To identify the sequential relationship, we exploit the fact that the international fragmentation of production only started to expand at a global scale in the 1990s, as a result of a dramatic fall in cross-border trade and communication costs. Domestic fragmentation, in turn, has been available to firms much longer. Therefore, initial

² See, among others, Jones (2011), Bartelme and Gorodnichenko (2015), Bernard et al. (2015), and Dhyne and Rubínová (2016).

³ See, for example, Jones and Kierzkowski (1990, 2001), Antràs (2003), Grossman and Helpman (2003), and Fally and Hillberry (2015).

⁴ See, for example, Antràs and Helpman (2004) and Grossman and Helpman (2005).

levels of DVC integration are unlikely to be affected by current levels of GVC integration.

The mechanism linking GVC and DVC integration that we focus on relates to the fragmentation costs associated with slicing up production, and to the costs of switching suppliers. We refer to fragmentation costs as all the costs paid by a firm deciding to source inputs externally, rather than producing them in-house.⁵ Here, we focus on that part of fragmentation costs that can be considered as one-off fixed costs. Such fixed costs entail, among other things, the necessary codifying of tacit knowledge, the downsizing of plants and workforce, the protection of intellectual property, the adaptation of the firm's structure, and learning about the fragmentation process.⁶ Once these costs have been paid for domestic fragmentation, they are not incurred again. This implies that DVC integration should serve as a stepping stone to subsequent GVC integration, particularly in industries characterized by relatively large fragmentation costs. Other things being equal, this would be reflected in a positive relationship between DVC and GVC integration.

Conversely, we refer to switching costs as the costs paid by a firm when it decides to stop sourcing an input from a supplier and starts sourcing the same input from another supplier, either located in the same country or abroad. Examples are contract termination fees and supplier searching and training costs. For a given level of domestic fragmentation, high switching costs should reduce subsequent international fragmentation making DVCs a stumbling block to GVC integration. Other things being equal, this would be reflected in a negative relationship between DVC and GVC integration.

Thus, in the presence of both fragmentation costs and switching costs, the sign of the relationship between DVCs and GVCs is dependent on the relative importance of switching vis-à-vis fragmentation costs and needs to be determined empirically. To that effect, we propose a novel measure of DVCs, equal to the share of domestically sourced inputs in domestic output, where

⁵ Fragmentation costs are incurred both if the sourcing occurs within national borders and if the sourcing occurs internationally.

⁶ This is analogous to the literature on learning to export where firms start their internationalization by exporting small amounts to easily accessible markets to learn about the export process and their export potential before expanding their activities (see for instance Schmeiser, 2012).

the latter is computed excluding foreign sourced inputs. We are able to compute this measure for up to 61 countries and 34 industries.⁷ We show that higher levels of domestic fragmentation at the beginning of our sample (mid 1990s) positively affect GVC integration, as measured by foreign value added in exports, at the end of the sample (late 2000s). In the benchmark specification, a one unit increase in DVC integration raises subsequent GVC integration by 2.4%. This suggests that DVC integration is a stepping stone for GVC integration, and that, overall, the fragmentation cost channel dominates the switching cost channel. Moreover, by taking DVCs into account, we are able to reduce the unexplained variation in GVC integration by about 30%.

To investigate the hypothesized role of fragmentation and switching costs for the relationship between DVCs and GVCs, we exploit manufacturing industry-level heterogeneity along two dimensions: the degree of product differentiation and contract intensity (i.e. relationship specificity). Fixed fragmentation costs are likely to be higher in industries producing a high share of differentiated goods since these industries tend to be more innovation- and skill-intensive, therefore they have higher costs of codifying tacit knowledge or downsizing as shown by Voigtländer (2014). Switching suppliers, in turn, is costlier in industries characterized by a high degree of contract intensity because it involves the payment of contract termination fees and similar costs that tend to increase with the relationship-specificity of inputs. Accordingly, we use Rauch (1999)'s classification for differentiated goods to proxy for fragmentation costs, and Nunn (2007)'s contract intensity measure to proxy for switching costs.⁸ The 26 manufacturing industries are then grouped into three categories: industries with high fragmentation and high switching costs, industries with low fragmentation and low switching costs, and industries with low fragmentation and high switching costs.⁹ In line with the suggested mechanism, we find that the positive role of DVCs for subsequent GVC integration is driven by the first two groups of industries, suggesting that fragmentation costs tend to be more relevant than switch-

⁷ A complete list of countries and industries is available in Appendix A, respectively in Tables A-1 and A-2.

⁸ Since Nunn (2007)'s contract intensity measure employs Rauch (1999)'s classification, we show in the robustness section that our results hold when alternative proxies for switching and fragmentation costs are used.

⁹ The case of high fragmentation and low switching costs is not observed in the data.

ing costs. For the third group of industries, which is characterized by high switching costs and low fragmentation costs, the positive effect of DVCs disappears entirely.

The remainder of the paper is organized as follows. Section 2 provides details on the measures employed to proxy for DVC and GVC, describes the data used and the identification strategy. Section 3 presents the empirical results. Section 4 concludes and suggests some policy implications.

2 Empirical methodology

In this section we discuss the methodology to examine the role of DVCs for subsequently linking into GVCs. To do so, we postulate the following reduced form model at the country-industry level:

$$gvc_{ikt} = \beta dvc_{iks} + \boldsymbol{\lambda}' \boldsymbol{\iota} + \boldsymbol{\gamma}' \mathbf{x}_{ikt} + \varepsilon_{ikt}, \quad (2.1)$$

where i indexes industries; k indexes countries; t and $s < t$ index time; the scalar β is the coefficient of interest, to be estimated; $\boldsymbol{\gamma}$ is a column vector of coefficients to be estimated; \mathbf{x} is a column vector of controls; $\boldsymbol{\lambda}$ is a column vector of fixed effects; $\boldsymbol{\iota}$ is a column vector of 1's; and ε is a random error term. The variables in \mathbf{x} , and the elements in and dimension of $\boldsymbol{\lambda}$, are specified in Section 2.3, where the identification strategy is discussed.

While equation (2.1) can give the net effect of DVCs for GVC integration, it does not provide information on the mechanism driving the relationship. Therefore, in a second step we proceed estimating a variant of equation (2.1) that exploits the industry-level variation in the data to analyze the suggested role of fragmentation and switching costs. For this purpose, we build proxies for industries' fragmentation and switching costs and then classify industries based on the relative magnitude of the costs into three groups: industries with high fragmentation and high switching costs (*HFHS* industries), industries with low fragmentation and low switching costs (*LFLS* industries), and industries with low fragmentation and high switching costs (*LFHS*

industries). We postulate the following reduced form model with interactions between DVCs and these industry-level characteristics:

$$gvc_{ikt} = \delta' (dvc_{iks} \times \mathbf{c}'_i) + \lambda' \iota + \gamma' \mathbf{x}_{ikt} + \varepsilon_{ikt}, \quad (2.2)$$

where δ is a column vector of coefficients to be estimated and $\mathbf{c}_i \equiv (HFHS_i \ LFLS_i \ LFHS_i)$.¹⁰

2.1 Measuring value chain integration and industry level characteristics

The literature has developed several measures for GVC participation that can be split up into backward and forward linkages. The former are related to intermediates sourced from abroad. The latter are related to domestic intermediates exported abroad. Since our channel between DVCs and GVCs relates to the costs of sourcing intermediate goods, we focus on backward linkages measures. These measures are typically based on Inter Country Input Output (ICIO) tables, which are $jl \times ik$ matrices that represent supply and demand relationships within and across industries (i, j) and countries (k, l) . A generic element m_{lk}^{ji} of an ICIO matrix gives the value of intermediate goods supplied by industry j of country l to industry i of country k . ICIOs indicate not only cross country linkages but also within country linkages. This implies that the same data source can be used to build both our DVC and GVC integration indicators, limiting the scope for measurement error. In the next subsections, we first discuss the standard measures of GVC integration proposed in the literature and also used in this study. Next, we move to our novel measure of DVC integration.

2.1.1 GVC integration

The recent literature on GVCs has used various indicators to capture the rise of global production networks. Recently, Hummels et al. (2001)'s Vertical Specialization measure and its refinements by Wang et al. (2013) and Koopman et al. (2014) have emerged as standard indicators, also

¹⁰ The \mathbf{c} vector does not contain a *HFLS* dummy since, as mentioned in footnote 9, the case of high fragmentation and low switching costs is not observed.

adopted in the present analysis.

The baseline GVC indicator is *fvax*, the foreign value added content in the production of exports. To construct this indicator, value added needs to cross at least two borders to be counted towards GVC trade. To obtain the required value added flows, it is necessary to decompose gross exports using information from ICIOs.¹¹ In the simplest case with two countries and two industries, the decomposition can be illustrated as follows:

$$V(I - A)^{-1}E = \begin{pmatrix} vae_{kk}^{ii} & vae_{kk}^{ij} & vae_{kl}^{ii} & vae_{kl}^{ij} \\ vae_{kk}^{ji} & vae_{kk}^{jj} & vae_{kl}^{ji} & vae_{kl}^{jj} \\ vae_{lk}^{ii} & vae_{lk}^{ij} & vae_{ll}^{ii} & vae_{ll}^{ij} \\ vae_{lk}^{ji} & vae_{lk}^{jj} & vae_{ll}^{ji} & vae_{ll}^{jj} \end{pmatrix},$$

where E is a 4×4 matrix in which the diagonal elements give each industry's gross exports for N industries and G countries, V is a 4×4 matrix in which the diagonal elements give each industry's value added to output ratio, and A is the 4×4 Input Output coefficient matrix, i.e. each element of A gives the intermediates that each industry supplies for \$1 of output in every other industry.¹² The intuition behind the decomposition becomes clear by recalling that these intermediate flows are the output of other industries, which means that they consist of value added and other intermediates themselves. $(I - A)^{-1}$ approximates these indirect links between industries and when combined with the values in the V matrix gives the actual value added flows so that the elements of the *vae* matrix are estimates of the industry-level value added origins of each industry's exports.¹³

Our benchmark GVC indicator *fvax* for industry i in country k is then given by:

$$fvax_{ik} \equiv \ln(FVAX_{ik}) = \ln\left(\sum_l \sum_j vae_{lk}^{ji}\right) \quad (2.3)$$

¹¹ This is technically implemented using the R package *decompr* by Quast and Kummritz (2015).

¹² In the general case of G countries and N industries, the dimension of the E , V , A (and, consequently, *vae*) matrices is $GN \times GN$.

¹³ A more exhaustive explanation of the approach can be found in Wang et al. (2013).

where $l \neq k$. Thus, fax_{ik} is equal to the sum of value added from all industries j of all foreign countries l in the exports of industry i in country k .¹⁴

As a simple alternative indicator, used in robustness exercises, we additionally calculate the log of imported inputs, $i2p$, following Baldwin and Lopez Gonzalez (2015). This measure implies a broader definition of GVC and can easily be derived from ICIOs as follows:

$$i2p_{ik} \equiv \ln(I2P_{ik}) = \ln\left(\sum_l \sum_j m_{lk}^{ji}\right), \quad (2.4)$$

where $l \neq k$.¹⁵

2.1.2 DVC integration

For DVC integration (the explanatory variable of interest in equation (2.1)) we suggest a novel indicator, applying the following approach. First, we decompose an industry's output into three parts based on the different inputs required to produce it: i) its own value added; ii) domestically sourced intermediate goods; and iii) foreign sourced intermediate goods. We refer to the combination of i) and ii) as 'virtual domestic output', since it is the virtual part of an industry's output that is produced with domestic content only. With this distinction at hand, we proceed to define our DVC indicator as the share of domestically sourced intermediate goods in virtual domestic output:

$$dvc_{ik} \equiv \frac{\sum_j m_{kk}^{ji}}{\sum_j m_{kk}^{ji} + va_{ik}}. \quad (2.5)$$

In other words, dvc is the share of domestic inputs in total output excluding foreign inputs. That is, it captures how much of an establishment's domestically produced output is supplied by other establishments instead of being produced in-house.¹⁶

The reason for calculating the DVC indicator in this particular way is as follows. We only

¹⁴ Sourcing from ISIC Rev. 3 group C (mining industry) is excluded to avoid effects stemming from price variations in commodities. In Section 3.3 we use further strategies to deal with the mining industry.

¹⁵ In this case, too, sourcing from ISIC Rev. 3 group C (mining industry) is excluded.

¹⁶ Note that, as for GVC measures, we exclude intermediates sourced from the mining industry (ISIC Rev. 3 group C).

look at virtual domestic output instead of total output to avoid a mechanical correlation with the dependent variable which, as discussed in Section 2.1.1 above, is a subset of foreign sourced intermediate goods. These, in turn, are a subset of total output (recall that we decompose total output into three parts: own value added, domestically sourced intermediate goods, and foreign sourced intermediate goods). Hence, using total output as denominator of our *dvc* measure would correlate dependent and independent variable. By considering only virtual domestic content, we break this mechanical link.¹⁷ An alternative solution to the problem of mechanical correlation between the main explanatory variable (*dvc*) and the dependent variable (*gvc*) would be the use of levels, rather than shares, of the *dvc* variable. This approach, however, would also create a positive (negative) mechanical correlation between the two variables in the presence of any complementarity (substitutability) between foreign and domestic inputs in production.¹⁸

To summarize, we avoid a mechanical correlation between our GVC and DVC indicators by constructing the former in levels and the latter in shares, using in its denominator only virtual domestic content. Table 1 shows the top and bottom ten countries in terms of the *dvc* variable in 1995 (averaged across all industries for each country).¹⁹

< Table 1 about here >

2.1.3 Industry fragmentation costs and switching costs

As benchmark proxy for fragmentation costs, we use Rauch (1999)'s classification for differentiated goods concurred to our industry-level data.²⁰ The reason for this choice is that industries

¹⁷ The difference between total output and virtual domestic output in 1995 was in any case minimal and, therefore, as shown in Section 3.3, the results are not sensitive to the choice of denominator of the *dvc* variable.

¹⁸ A number of recent papers that measure gains from trade in the presence of intermediates consider the relationship between foreign input use and domestic input use. Blaum et al. (2015) combine domestic and foreign inputs in a constant elasticity of substitution (CES) function, with elasticity of substitution ε ranging between 1.4 and 2.4. Similarly, in Tintelnot et al. (2017), firms use a CES input bundle of workers and domestic and foreign inputs, with elasticity of substitution $\rho > 1$. The authors use values of ρ ranging from 1.5 to 2. Antràs et al. (2017) assume that, from the perspective of firms, inputs are differentiated by country of origin, with an elasticity of substitution across inputs from any two locations equal to $1 + \theta$. Their benchmark estimate of θ – which corresponds to the shape parameter of the Fréchet distribution of firm-specific intermediate input efficiencies – is equal to 1.789.

¹⁹ A complete list by country and industry is available in the Appendix. See Tables A-3 and A-4.

²⁰ The Rauch classification treats goods as differentiated if they are neither reference priced nor traded on organized exchanges.

with a high share of differentiated goods tend to be more innovation and skill intensive.²¹ As a result, we can assume that codifying tacit knowledge, downsizing, and other fragmentation related activities are more costly for these industries.

Regarding switching costs, Nunn (2007)'s contract intensity (or relationship specificity) measure is a straightforward proxy. Different industries require different amounts of relationship specific inputs. When these inputs are supplied at arm's length, complex contracts are needed to address the resulting hold up problem. Nunn (2007) shows that, as a result, industries differ in their contract intensity. A substantive share of fixed switching costs, such as contract cancellation fees, only occur when intermediates cannot be sourced from organized markets, but are sourced using contract-based relations with suppliers. Nunn (2007)'s indicator measures the share of intermediates that are sourced through such contract-based relations in total intermediates, and is thus a natural proxy for our exercise.²²

Table 2 lists the industries separated into the three categories and shows that the selected proxies create a sensible allocation. *LFLS* industries like 'Food and beverages' typically source inputs from organized markets and, thus, should find it easy to switch between suppliers. At the same time, their production tends to be neither skill nor innovation intensive and so fragmentation should be equally simple. In contrast, the identified *HFHS* industries such as 'Motor vehicles' or 'Electronics' are dependent on a highly skilled workforce and complex intermediates. Lastly, *LFHS* industries such as 'Electrical machinery' (e.g. cables) require complex intermediates but production is homogeneous and non complex.

< Table 2 about here >

²¹ See, for example, Voigtländer (2014) for evidence on this relationship.

²² A detailed explanation of the classifications of industries is available in Appendix B. Note that Nunn uses the Rauch classification to determine if an intermediate is sourced on a contractual basis. Thus, our switching and fragmentation cost proxies depend indirectly and directly on the Rauch classification. However, since the Nunn classification looks at an industry's intermediates and the Rauch classification at an industry's output, there is no strong mechanical correlation between the two. We show in Section 3.3 that our results are robust to employing alternative proxies for switching and fragmentation costs such as Costinot (2009)'s complexity classification.

2.2 Descriptive evidence

When taking our GVC and DVC indicators to the data, we find preliminary evidence that DVC integration is a stepping stone for GVC integration. Figure 2 unconditionally correlates GVCs in 2008 (in logs) with the measure of DVCs in 1995. The slope coefficient is equal to 0.47. The stepping-stone effect of DVC integration for GVC integrations also constitutes suggestive evidence that the fragmentation cost channel dominates the switching cost channel.

< Figure 2 about here >

Similarly, when we look at our DVC integration indicator by country (Table 1), we see that countries that are well integrated into GVCs also had strong domestic linkages in 1995. Interestingly, however, while the top ten is dominated by countries traditionally highly involved into GVCs, such as China, Slovakia, or Korea, it also contains countries like Croatia and New Zealand, which do not exhibit strong linkages into GVCs. The same pattern holds when examining the bottom ten. While Cambodia and Greece lag behind in GVC participation, Mexico is highly integrated. This suggests heterogeneity in the effects of DVCs, which will be taken into account in the empirical analysis.

2.3 Identification

Estimation results of the reduced form equations (2.1) and (2.2) could in particular be affected by omitted variable bias. For instance, high levels of GVC integration could be driven by a country's institutions, which might also affect domestic production fragmentation. To deal with this concern, we use a combination of fixed effects and control variables that account for structural and policy determinants of GVC participation identified in the literature. More specifically, we use industry fixed effects to control for structural differences across industries that might correlate at once with GVC and DVC integration. Heavy manufactures and downstream industries, for instance, are more likely to develop backward linkages than agriculture. In addition,

we include country-broad industry fixed effects to take care of (non time-varying) policy and structural differences across countries – in terms of country size, economic development, remoteness, institutions, endowments as well as the industrial structure of the country – that may vary across broadly defined industries in each country.²³

Industry and country-broad industry fixed effects fail to control for any potentially omitted variable that varies across countries-industries and might be correlated with DVC integration, even within broad industries. We include two such variables: i) a dummy equal to one for comparative advantage industries, *ca*; and ii) industry level ‘composite’ import tariffs, *ctau*. The comparative advantage dummy synthetically accounts for all those determinants of patterns of specialization of a given country in a given industry (e.g., country labour abundance in a labour-intensive industry). It is calculated using Balassa (1965)’s revealed comparative advantage index, but on the basis of value added exports.²⁴ The ‘composite’ import tariffs variable captures differences in the trade policy environment at the country-industry level. For each industry in a given country, it is calculated as the weighted average of tariffs that the inputs into the industry face, with weights being given by the input output coefficient. We follow the literature and use United States’ input output coefficients under the assumption that they are determined by technological factors as opposed to market distortions.²⁵

Considering that the aim is to examine how initial patterns in DVCs have affected subsequent GVC integration, we use *dvc* values for 1995 and *fvax* values for 2008. This also addresses potential concerns about a reverse causality bias since the rapid expansion of GVCs only started around 1995 (Wang et al., 2017 show that in 1995 GVCs only accounted for 3% of global production). This limits any potential reverse effect on DVC integration values in 1995.

²³ As the dependent variable is at the country-industry level, we cannot use country-industry fixed effects to control for a country’s industrial structure. To be nevertheless as conservative as possible, we form country-broad industry fixed effects by interacting country fixed effects with fixed effects for the six ‘broad industries’ defined in the last column of Table A-2. These country-broad industry fixed effects control for the fact that larger industries within a country have more scope for fragmentation.

²⁴ The Balassa index is given by $\frac{e_{ik}/e_k}{\sum_k e_{ik}/\sum_k e_k}$. It compares an industry’s share in a country’s exports, e_k , to that industry’s share in global exports. An index value larger than 1 indicates comparative advantage industries.

²⁵ See, for instance, Hsieh and Klenow (2009).

The benchmark reduced form model for the effect of DVCs on GVC integration (equation (2.1)) is then estimated as:

$$gvc_{ik2008} = \beta dvc_{ik1995} + \gamma_1 ca_{ik2008} + \gamma_2 ct\tau_{ik2008} + \lambda + \lambda_i + \lambda_{bk} + \varepsilon_{ik2008}, \quad (2.6)$$

where β is the coefficient of interest; λ is a constant; and λ_i and λ_{bk} are respectively industry and broad industry-country fixed effects. Since the mechanism determining the role of DVCs for GVC integration is ambiguous as outlined in the introduction, we do not have a prior on the sign of β .

The benchmark reduced form model for testing the mechanism behind the DVC-GVC link (equation (2.2)), in turn, is estimated as:

$$gvc_{ik2008} = \delta_1 dvc_{ik1995} \times LFLS_i + \delta_2 dvc_{ik1995} \times HFHS_i + \delta_3 dvc_{ik1995} \times LFHS_i + \gamma_1 ca_{ik2008} + \gamma_2 ct\tau_{ik2008} + \lambda + \lambda_i + \lambda_{bk} + \varepsilon_{ik2008}. \quad (2.7)$$

In equation (2.7), δ_1 gives the effect of DVCs on subsequent GVC integration in *LFLS* industries; δ_2 gives the effect for *HFHS* industries; and δ_3 gives the effect for *LFHS* industries.

A positive δ_1 would suggest that fragmentation costs play a larger role than switching costs, while a negative δ_1 would imply the opposite. We include the *HFHS* dummy as falsification test, since we expect δ_2 not to be significantly different from δ_1 since in both industry categories the relative level of fragmentation vis-à-vis switching costs is identical. In other words, if the channel we propose is at work, there is no reason to believe that the effect of DVCs on GVCs should be different in *HFHS* industries as compared to *LFLS* industries. The coefficient of interest in equation (2.7) is δ_3 , which is expected to be significantly smaller than δ_1 and δ_2 . This is because in *LFHS* industries the relative level of switching costs, as compared to fragmentation costs, is higher than in the benchmark *LFLS* industries. Accordingly, DVC integration in these industries is less likely to lead to GVC integration.

The proposed strategy of adding industry and broad-industry country fixed effects, as well as the *ca* and *ctau* controls, does not guarantee that the empirical correlation between *dvc* (alone or interacted with industry fragmentation/switching costs dummies) and *gvc* is free of any omitted variable bias. Therefore, as described in Section 3.3, panel estimations using data for the years 2000 and 2005 in addition to 1995 and 2008 are presented. These estimations allow to include a rich structure of fixed effects, namely country-year, industry-year, and industry-country, controlling for any unobserved factor that might be correlated with GVCs and DVCs and varies across two of the three dimensions (countries, industries and time) of the data. This should minimize concerns related to omitted variable bias.

Moreover, the approach of using 1995 values for the *dvc* variable and 2008 values for the *gvc* variable does not automatically exclude any reverse causality. As also detailed in Section 3.3, we thus estimate regressions with an interaction term between *gvc* and a high income dummy as a way of addressing these concerns. The reason is that if there is reverse causality flowing from GVCs to DVCs, we would like our *dvc* values to be from a point in time where the extent of GVCs was negligible, so that the estimates would be unbiased even in the presence of reverse causality. In 1995 the extent of GVC integration might not have been negligible in developed countries, but was certainly so in many developing countries.

Killing two birds (omitted variables and reverse causality) with one stone, we further propose an instrumental variable (IV) strategy that exploits exogenous variation in the DVC variable. To obtain such exogenous variation, we regress *dvc* in the first stage on the weighted average of other countries' *dvc* values, with the weights given by country size (as proxied by GDP).²⁶ The intuition is the following. Think of DVC values as having three components: a country-specific, an industry-specific and a country-industry component. The first two components are absorbed by fixed effects, and are not a source of concern. The third component itself can further be decomposed in two parts: i) an idiosyncratic part specific to each country-industry (e.g. man-

²⁶ The instrument for dvc_{ikt} is, therefore, the variable $dvc_{ikt}^{IV} \equiv \sum_{l \neq k} dvc_{ilt} \times gdp_{lt}$, where gdp_{lt} is the share of country l in total GDP of all countries in the sample at time t . The weights sum up to one.

agerial ability, reputation); and ii) a part correlated across countries due to the interaction of structural industry characteristics and country characteristics (e.g. market size matters more for domestic fragmentation in complex industries than in non-complex industries). The idiosyncratic part might introduce an omitted variable bias – for instance if high managerial ability specific to an industry in a country might lead to both domestic and foreign fragmentation. In contrast, the correlated part should be free of such bias, because the individual contribution of the industry- and country-specific components is controlled for by fixed effects, so that identification is obtained only through the exogenous interaction of these variables. By using other countries' DVC variables as an instrument, the estimates are only identified by this exogenous part, which is by definition uncorrelated with the endogenous idiosyncratic part. Therefore, the average value of *dvc* in other countries is a legitimate instrument for *dvc*, since it both meets the relevance and the exclusion requirements. We additionally weight the instrument by country size since we assume that the DVC values of larger countries in our sample are closest to a true optimal undistorted DVC value because large countries have a larger set of available suppliers, which means that their DVC values are not artificially restricted by the absence of adequate domestic outsourcing opportunities.

2.4 Data

For the calculation of the DVC and GVC indicators, as well as the revealed comparative advantage measure, we employ the most recent and most advanced release of OECD ICIO tables, covering 61 countries and 34 2 digit ISIC Rev. 3 industries for the years 1995, 2000, 2005, and 2008 to 2011. To create ICIOs, the OECD combines national IO tables with international trade data. As OECD countries have a harmonized construction methodology, potential discrepancies between national IO tables should be minor. Furthermore, the advanced harmonization across countries reduces to a minimum the use of proportionality assumptions to derive the ratio of imported intermediates in an industry's demand. The OECD has used elaborate techniques

to deal with processing trade. Due to the outstanding role of processing trade in GVCs, this implies a significant improvement for the reliability of the database.²⁷

For the tariff measure, we additionally take HS 2 digit MFN tariffs from the UNCTAD TRAINS database and concord them to the ICIOs' ISIC Rev. 3 classification employing a concordance table provided by WITS. The Rauch classification is available at the 4 digit SITC Rev. 2 level. To concord it to ISIC Rev. 3, we need to apply a crosswalk via SITC Rev. 3. The necessary concordance tables are provided by Eurostat. In cases where ISIC industries are concorded to both homogeneous and differentiated SITC commodities, we use trade weights from COMTRADE to determine if an ISIC industry is differentiated or homogeneous.²⁸

For additional specifications we employ as supplementary controls the distance of country k to the closest GVC hub (China, Germany, Japan, and the United States) from CEPII and GDP data from the World Development Indicators. We also perform several robustness checks, some of which necessitate additional data. Firstly, in one robustness exercise we replace the Rauch classification for fragmentation costs with Costinot (2009)'s measure of industrial complexity. It assesses how many days of training an average worker in a given industry needs to complete a job, and is hence a direct measure of an industry's skill content. The data is available for a subset of 3 digit SIC codes at the 1972 revision. To match these with ISIC codes, we first concord them to the 1987 SIC revision using concordance tables by the NBER. We then use a concordance table between SIC87 and ISIC Rev. 3 built by Statistics Canada.²⁹ Secondly, in a separate exercise we replace the Nunn's proxy for contract intensity with US capital intensity data from the OECD. In capital intensive industries, suppliers are often required to make large sunk investments. Consequently, they are likely to demand contracts with high cancellation fees. This makes capital intensity a likely proxy for switching costs. This data is readily available at the ISIC level.

²⁷ See Koopman et al. (2012) for an analysis of China's processing trade.

²⁸ See Appendix B for details.

²⁹ Details are provided in Appendix B. All concordances and classifications are available from the authors upon request.

The country coverage comprises developing and developed economies, allowing to examine whether the level of development plays any role in shaping the relationship between domestic and global value chains. In addition, the industry coverage includes primary, manufacturing, and services sectors. The latter play an increasingly dominant role in value chains. However, there are no tariffs nor Rauch measures available for services. Therefore, we use estimates based only on the manufacturing and primary sectors as benchmark but perform robustness checks that include services without the tariff control, assuming that all services industries are differentiated and excluding non-tradeable services industries such as education, health, or construction. Concerning the time period examined, we use 1995 and 2008 because these are the earliest and latest available data points before the start of the prolonged global financial crisis. Finally, we exclude in the benchmark estimations the oil exporting countries Saudi Arabia and Brunei; Chinese Taipei, for which no tariff data is available; and the United States, to avoid endogeneity stemming from the construction of the weighted tariff and capital intensity variables.³⁰

3 Results

This section is structured in three parts. We start by discussing the outcome of equation (2.6), which estimates the net effect that higher values of DVC integration have for subsequent GVC integration. In the second part, we assess the mechanism behind this effect by disentangling the role of fragmentation and switching costs as identified by equation (2.7). In the last part of the section we present a series of robustness checks.

3.1 The net effect of DVCs on GVCs

Table 3 reports the results for the first part of this exercise. The coefficient on *dvc* is positive and statistically significant in all specifications, indicating that domestic value chains facilitate

³⁰ Exclusion of the United States from the sample is standard practice in the literature that uses US IO weights – see for instance Beverelli et al. (2017).

subsequent GVC integration.

< Table 3 about here >

Columns (1) and (2) give the results for primary and manufacturing sectors only, while columns (3) and (4) include industries in the services sector, but at the cost of not controlling for tariffs. While all estimates are similar in magnitude and significance, we take the point estimate in column (2) as benchmark because it includes the full set of fixed effects described in equation (2.6) and it is comparable to the analysis in Section 3.2.

The benchmark specification suggests that a one unit increase in initial DVC integration leads to 2.4% higher subsequent GVC integration. This implies that a one standard deviation increase in DVC integration raises GVC integration by approximately 0.2 standard deviations or 0.4%. In addition, back of the envelope calculations comparing R squared of the benchmark regression with and without the *dvc* measure suggest that DVCs can reduce the unexplained variation in contemporaneous GVC integration by about 30%.

Since our benchmark estimates only allow for within-broad-industry-country variation, they might hide gains that stem from variation in DVC integration across countries which are emphasized by some studies such as Bartelme and Gorodnichenko (2015). Therefore, we additionally report in columns (1) and (3) results that only control for industry fixed effects, including in the vector of controls the following three variables: log of per capita GDP at constant 2005 prices; log of GDP at constant 2005 prices; and log distance to the closest GVC hub (CHN, DEU, JPN, or USA). These variables capture basic country characteristics such as size, development status, and remoteness. Since the estimates in columns (1) and (3) are potentially subject to an omitted variable bias due to the exclusion of country fixed effects, we interpret them as upper bounds of the DVC effect. Table 3 shows that the coefficients are indeed larger in magnitude than the benchmark estimates, although the difference is not large.

In summary, these findings highlight a further advantage of comprehensive domestic linkages through their contribution to GVC participation. In the light of the literature that has unveiled productivity enhancing effects of GVC integration (Kummritz, 2016), our results can help explain why efficient DVCs facilitate economic development. Our findings also speak to the literature on the determinants of GVC integration. We show that differences in pre-existing domestic linkages can account for differences in observed GVC linkages.

The descriptive evidence discussed in Section 2.1 has however illustrated that there are countries with strong DVC linkages which are not well integrated into GVCs. To unveil the heterogeneity in the GVC-DVC relationship, we now empirically assess the relevance of the mechanism related to the industry-specific costs discussed above.

3.2 The role of fragmentation and switching costs

In this section, we estimate equation (2.7), which separates the effects of DVCs on GVCs in *LFHS* (low fragmentation, high switching costs) industries as opposed to *HFHS* (high fragmentation, high switching costs) industries, and *LFLS* (low fragmentation, low switching costs) industries.

As a reminder, we expect the coefficient on the *LFHS* dummy to be significantly smaller than the coefficients on the *HFHS* and the *LFLS* dummies. This is due to the fact that the former dummy captures industries in which switching costs tend to be high, which makes a change from domestic to foreign suppliers costly, but fragmentation costs low, which shuts down the stepping stone channel of DVCs for GVCs. In addition, we expect the coefficients on the *HFHS* and the *LFLS* dummies to be not significantly different, since the relative incidence of fragmentation and switching costs should be identical in the two cases.

Table 4 reports the corresponding results. As in Table 3, columns (1) and (2) exclude industries in the services sector, while columns (3) and (4) include them. We replicate in columns (1) and (3) the basic results presented in Table 3 for comparison, while columns (2) and (4) report the new estimates from equation (2.7).

< Table 4 about here >

The results are in line with our expectations. The coefficients on the interaction of the *dvc* variable with the *HFHS* and the *LFLS* dummies (δ_1 and δ_2 in equation (2.7)) are positive and statistically significant in both samples, while the coefficients on the interaction with the *LFHS* dummy (δ_3 in equation (2.7)) is smaller and, in fact, not statistically different from zero. This is strong evidence in favour of the proposed channel which links initial DVC patterns to subsequent GVC integration by taking into account the varying importance of fragmentation and switching costs across industries.

Our preferred specification in column (2) suggests that in *HFHS* and *LFLS* industries, a one percentage point increase in initial DVC linkages increases contemporaneous GVC integration by 2.3 to 3.1%. However, this positive effect of DVCs is not observed in *LFHS* industries.

In terms of our proposed channel, these findings are further evidence that fragmentation costs are more important than switching costs for the DVC-GVC link. This is because in industries where both costs are high (low), the positive effect of DVCs prevails. Similarly, even when switching costs are high but fragmentation costs are low, the net effect is around zero instead of turning negative. Thus, the net effect that we found in Section 3.1 is confirmed and can be explained by the fact that fragmentation costs are more relevant than switching costs for moving from DVCs to GVCs.

These findings can also explain the observed heterogeneity in the DVC-GVC link as presented in the introduction. In Figure 3 we revisit the matched country pairs from Figure 1. There, we grouped countries based on the known structural and policy determinants of GVC integration and showed that there remains substantial unexplained variation in the data. Here, we add leading export industries to the chart. It is striking that all the leading export industries of the under-performers (respectively, ‘Paper – PAP’ for Sweden, and ‘Chemicals – CHM’ for Switzerland and for Romania) fall into the *LFHS* category whereas the comparison group is more specialized in *HFHS* industries (‘Computer, electronic and optical products – CEQ’ in the

case of Singapore, ‘Textiles, leather and footwear – TEX’ in the case of Bulgaria).³¹ This means that in these three under-performing countries the structure of the economy causes domestic linkages to prevent GVC integration, despite the fact that other factors, such as location, would facilitate GVC integration. For policy makers this implies that they need to take into account both the strength of domestic linkages and the structure of their economies when they develop GVC strategies.

< Figure 3 about here >

3.3 Robustness tests and IV regressions

We expose our results to a battery of robustness checks starting with the identification assumptions. While we have argued that in 1995 reverse causality from GVCs to DVCs should have been minor, since only 3% of global production was related to GVCs, there is evidence that in high income countries GVCs did already play a role. We augment equation (2.6) with an interaction term between the DVC measure and a high income dummy, equal to one if a country was classified as high income in the World Bank classification in 1995. This allows us to see if the effect is also present in developing economies, for which the assumption of the absence of reverse causality is not questionable. This estimation can also provide insights on a possible development dimension to our findings. The results are in column (2) of Table 5.³² The high income interaction term is not significant. We interpret this result (which is robust to several alternative constructions of country groups) as an indication that the effect of DVC on GVC integration is not driven by reverse causality, since the effect is also present in developing economies, where GVCs linkages only developed in relatively recent years.

< Table 5 about here >

³¹ An exception is Belgium, which is also specialized in ‘Chemicals’. However, its DVC integration levels in 1995 were fairly low in this industry while Switzerland, Sweden and Romania had high values of DVC integration in this industry.

³² Column (1) of Table 5 reproduces the estimates of the preferred specification of column (2) of Table 3 for convenience.

Next, we further address concerns about omitted variable bias by fully exploiting the time variation in our data. More specifically, we additionally use data for the years 2000 and 2005 and estimate a panel version of equation (2.6). The specification includes a one period lagged DVC variable (following Bartelme and Gorodnichenko, 2015) and country-year, industry-year, and industry-country fixed effects. The specification without the time dimension remains our preferred model since the objective is to assess the effect of initial DVC conditions unaffected by the rise of GVCs. However, the advantage of this robustness check is that we can add a rich structure of fixed effects, controlling for any unobserved factor that might be correlated with GVC and DVC and varies across two of the three dimensions (countries, industries and time) of the data. Column (3) of Table 5 shows that the coefficient on *dvc* remains positive and statistically significant. Its magnitude decreases by 75%. This can be explained by a potential linear accumulation of the effect over time.

The results of the instrumental variable regression using the variable dvc^{IV} described in Section 2.3 as an instrument for *dvc* are in column (4) of Table 5. All coefficients are in line, in terms of sign, magnitude and statistical significance, with the benchmark coefficients reported in column (1). The Cragg-Donald Wald F statistic of 176.6 is well above the corresponding critical values tabulated in Stock and Yogo (2005), suggesting that the instrument is relevant. Note further that the first-stage correlation between the instrument and *dvc* is statistically significant. Taken together, the results of columns (1)-(3) of Table 5 allow to conclude that domestic linkages causally affect GVC participation, in other words they are stepping stones for integration in global value chains.

The stepping-stone result is subject to five further robustness checks addressing measurement and sample selection issues. The results of these tests are displayed in columns (5)-(9) of Table 5. First, we test if our results are dependent on the selected GVC measure. For this purpose, we replace *fvax* with the *i2p* indicator of Baldwin and Lopez Gonzalez (2015), which is simply the value of intermediate imports (see equation 2.4 and Section 2.1.1). We find in

column (5) of Table 5 that this variation has no relevant impact on our estimates.

We then proceed to vary the sample composition by including the previously excluded countries Saudi Arabia, Brunei, and the United States and, in a separate exercise, non tradeable services industries to examine if the results are sensitive to specific sample changes (see Table A-2 for the list of non-tradeable industries). Columns (6) and (7) of Table 5 show that the results are not affected in any meaningful way.

In a next step, we replace the GVC value of 2008 with the values of 2005 (in column (8)) or 2011 (in column (9)). This allows assessing whether the results are dependent on a specific time period, in particular whether the global financial crisis is relevant for the DVC-GVC relationship. The estimates in columns (8) and (9) of Table 5 suggest that this is not the case. In addition, the latter result that uses GVC values for 2011, combined with the estimates from the robustness exercise that uses a panel setup (column (3)) reveals suggestive evidence on the role of DVCs in the current slowdown in GVC expansion.³³ The panel estimates suggest that the effect of DVCs accumulates over time until 2008. However, the coefficient for 2008 is not statistically different from the coefficient for 2011, implying that the accumulation has ended. This would be in line with the fact that the boost that initial DVC integration provided to GVC expansion has been fully exploited which, in turn, would contribute to explaining the observed slowdown in GVC expansion.

We now turn to a set of robustness checks concerning the mechanism behind the DVC-GVC link of Table 4, starting with the results of IV regressions. The IV results, presented in column (2) of Table 6, are similar, in terms of sign, magnitude and statistical significance, to the benchmark coefficients of column (1).³⁴ Namely, the coefficients on the interaction of the *dvc* variable with the *HFHS* and the *LFLS* dummies are positive and statistically significant, and the coefficient on the interaction with the *LFHS* dummy is smaller and not statistically

³³ See, for instance, Degain et al. (2017) for recent evidence that the GVC expansion has stalled since 2008.

³⁴ The benchmark coefficients of column (1) of Table 4 are the ones of column (2) of Table 4, and are reported for convenience.

different from zero. The value of the Cragg-Donald Wald F statistic of 54.65 again suggests that the instrument is relevant, and the first-stage correlation between the instruments (interaction terms between dvc^{IV} and industry dummies) and the variables to be instrumented (interaction terms between dvc and industry dummies) – presented in columns (2.a, 2.b and 2.c) – are statistically significant. We therefore confirm that the mechanisms behind the causal effect of initial DVC patterns on subsequent GVC integration are related to the varying importance of fragmentation and switching costs across industries, as discussed in Section 3.

< Table 6 about here >

The next three columns of Table 6 present the results of three robustness checks that concern the measuring of industry dummies that proxy for fragmentation and switching costs.

Firstly, we vary the cutoff condition which classifies industries as differentiated and, indirectly, as contract intensive. For this, it is necessary to recall that the Rauch classification is at the 4 digit SITC level, while our industry classification used to construct the GVC and DVC variables is the 2 digit ISIC. As a results, several SITC industries are assigned to one ISIC industry. As explained in detail in Appendix B, in the benchmark estimation, we define a 2 digit ISIC industry as differentiated if the trade share of differentiated 4 digit SITC industries matched with the 2 digit ISIC industry exceeds the mean value across all ISIC industries. We now use the median as cutoff instead. This reduces our *LFHS* industries to ‘Pulp and paper products’, ‘Chemicals’, and ‘Mining and quarrying’, with the other two in the last column of Table 2 being now classified as *HFHS* industries.

Column (3) of Table 4 shows the estimates for this slightly stricter cutoff. Qualitatively, the results hold while the magnitude of the *LFHS* coefficient increases at the expense of the coefficient on the *HFHS* interaction. Since we increased the cutoff, the ratio of switching to fragmentation costs in *LFHS* industries has increased causing a reclassification of *LFHS* industries as *HFHS* industries. Therefore, the change in magnitudes is sensible and confirms our channel.

In the two next robustness exercises, we change the measures used to classify the industries into their respective categories. First, we use Costinot (2009)'s industrial complexity indicator to proxy for fragmentation costs. As discussed in Section 2.4, it is a proxy for an industry's average skill intensity and can accordingly capture the costs of downsizing and protecting intellectual property well. The results in column (4) of Table 6 show that this change has no relevant impact on our estimates. As a second variation, we use capital intensity as an alternative proxy for switching costs and equally find no meaningful impact on our results, as can be seen in column (5) of Table 6.

Overall, the two main results concerning the net effect of DVCs for GVC integration, and the mechanisms behind this link, are robust to variations in the identification assumptions, the sample composition, and to the employed indicators and proxies. Therefore, we conclude that DVCs and their interaction with a country's industrial structure are an important driving factor for integration in GVCs.

4 Conclusions

GVCs have become a dominant factor in international trade. Recent theoretical and empirical research has shown that integration in GVCs can lead to gains in productivity and welfare. This begs the question of what drives GVC participation. While previous research has revealed a set of important structural factors and policies, there remains substantial unexplained variation in GVC integration patterns across countries.

This paper has shed light on a new determinant of these patterns by providing evidence that initial levels of DVC integration affect contemporaneous GVC integration. In other words, levels of domestic fragmentation established before the rise of GVCs can explain a relevant share of current variation in GVC integration. The results hold for countries at varying stages of development, over varying time periods, and when changing the sample composition and identifying assumptions. In our preferred specification, a one standard deviation increase in

DVC integration raises subsequent GVC integration by 0.4%. Besides, we are able to decrease the unexplained variation in GVC integration by about 30%.

The mechanisms behind the DVC-GVC link investigated in this paper are related to industry-specific fixed costs of fragmentation and of switching suppliers. On the one hand, high fixed fragmentation costs allow, due their sunk nature, DVCs to act as stepping stones for GVCs. On the other hand, high fixed costs of switching suppliers should lead to a negative association between DVCs and GVCs. Our findings on the net effect suggest that the fragmentation cost channel is more important than the switching cost channel. Exploiting the industry variation in fragmentation and switching costs in the data, we have confirmed that this is indeed the case. The facilitating effect of DVCs is absent in industries characterized by low fragmentation costs, but high switching costs.

The policy implications are relevant and straightforward. Key barriers to GVC integration might stem from the interplay of domestic linkages and industrial structure. These linkages should be taken into account when designing and evaluating GVC integration strategies.

References

- Antràs, P., 2003, “Firms, Contracts, and Trade Structure,” *Quarterly Journal of Economics*, 118, 1375–1418.
- Antràs, P., T. C. Fort, and F. Tintelnot, 2017, “The Margins of Global Sourcing: Theory and Evidence from U.S. Firms,” *American Economic Review*, 107, 2514–64.
- Antràs, P., and E. Helpman, 2004, “Global Sourcing,” *Journal of Political Economy*, 112, 552–580.
- Balassa, B., 1965, “Trade Liberalisation and “Revealed” Comparative Advantage,” *The Manchester School*, 33, 99–123.
- Baldwin, R., and J. Lopez Gonzalez, 2015, “Supply-chain Trade: A Portrait of Global Patterns and Several Testable Hypotheses,” *The World Economy*, 38, 1682–1721.
- Baldwin, R., and F. Robert-Nicoud, 2014, “Trade-in-goods and trade-in-tasks: An integrating framework,” *Journal of International Economics*, 92, 51–62.
- Bartelme, D., and Y. Gorodnichenko, 2015, “Linkages and Economic Development,” NBER Working Paper No. 21251, National Bureau of Economic Research.
- Bernard, A. B., A. Moxnes, and Y. U. Saito, 2015, “Production networks, geography and firm performance,” NBER Working Paper No. 21082, National Bureau of Economic Research.
- Beverelli, C., M. Fiorini, and B. Hoekman, 2017, “Services trade policy and manufacturing productivity: The role of institutions,” *Journal of International Economics*, 104, 166–182.
- Blaum, J., C. Lelarge, and M. Peters, 2015, “The Gains From Trade in Firm-Based Models of Importing,” unpublished manuscript, available at https://www.brown.edu/Departments/Economics/Faculty/Joaquin_Blaum/GainsFromInputTrade.pdf.

Costinot, A., 2009, “On the origins of comparative advantage,” *Journal of International Economics*, 77, 255–264.

Degain, C., B. Meng, and Z. Wang, 2017, “Recent trends in global trade and global value chains,” in *Global Value Chain Development Report 2017*, World Bank, IDE-JETRO, OECD, RCGVC-UIBE, WTO.

Dhyne, E., and S. Rubínová, 2016, “The supplier network of exporters: Connecting the dots,” Working Paper No. 296, National Bank of Belgium.

Fally, T., and R. Hillberry, 2015, “A Coasian Model of International Production Chains,” NBER Working Paper No. 21520, National Bureau of Economic Research.

Grossman, G. M., and E. Helpman, 2003, “Outsourcing Versus FDI in Industry Equilibrium,” *Journal of the European Economic Association*, 1, 317–327.

Grossman, G. M., and E. Helpman, 2005, “Outsourcing in a Global Economy,” *Review of Economic Studies*, 72, 135–159.

Grossman, G. M., and E. Rossi-Hansberg, 2008, “Trading Tasks: A Simple Theory of Offshoring,” *American Economic Review*, 98, 1978–97.

Hirschman, A. O., 1958, *The strategy of economic development*, New Haven: Yale University Press.

Hsieh, C.-T., and P. J. Klenow, 2009, “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics*, 124, 1403–1448.

Hummels, D., J. Ishii, and K.-M. Yi, 2001, “The nature and growth of vertical specialization in world trade,” *Journal of International Economics*, 54, 75–96.

Johnson, R. C., and G. Noguera, 2012a, “Accounting for intermediates: Production sharing and trade in value added,” *Journal of International Economics*, 86, 224–236.

Johnson, R. C., and G. Noguera, 2012b, “Fragmentation and Trade in Value Added over Four Decades,” NBER Working Paper No. 18186, National Bureau of Economic Research.

Jones, C. I., 2011, “Intermediate Goods and Weak Links in the Theory of Economic Development,” *American Economic Journal: Macroeconomics*, 3, 1–28.

Jones, R. W., and H. Kierzkowski, 1990, “The role of services in production and international trade: a theoretical framework,” in R. W. Jones, and A.O. Krueger (ed.), *The Political Economy of International Trade: Essays in Honour of Robert A. Mundell*, MIT Press, Cambridge, MA.

Jones, R. W., and H. Kierzkowski, 2001, “A framework for fragmentation,” in S.W. Arndt, and H. Kierzkowski (ed.), *Fragmentation: New Production Patterns in the World Economy*, Oxford University Press, New York.

Koopman, R. B., Z. Wang, and S.-J. Wei, 2012, “Estimating domestic content in exports when processing trade is pervasive,” *Journal of Development Economics*, 99, 178–189.

Koopman, R. B., Z. Wang, and S.-J. Wei, 2014, “Tracing Value-Added and Double Counting in Gross Exports,” *American Economic Review*, 104, 459–94.

Kowalski, P., J. Lopez Gonzalez, A. Ragoussis, and C. Ugarte, 2015, “Participation of Developing Countries in Global Value Chains,” OECD Trade Policy Paper No. 179, Organisation for Economic Co-operation and Development.

Kummritz, V., 2016, “Do Global Value Chains Cause Industrial Development?,” CTEI Working Paper No. 2016-01, Centre for Trade and Economic Integration.

Nunn, N., 2007, “Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade,” *Quarterly Journal of Economics*, 122, 569–600.

Quast, B. A., and V. Kummritz, 2015, “*decompr*: Global Value Chain decomposition in R,” CTEI Working Paper No. 2015-01, Centre for Trade and Economic Integration.

Rauch, J. E., 1999, “Networks versus markets in international trade,” *Journal of International Economics*, 48, 7–35.

Schmeiser, K. N., 2012, “Learning to export: Export growth and the destination decision of firms,” *Journal of International Economics*, 87, 89–97.

Stock, J., and M. Yogo, 2005, “Testing for Weak Instruments in Linear IV Regression,” in Donald W. K. Andrews (ed.), *Identification and Inference for Econometric Models*, New York: Cambridge University Press.

Timmer, M. P., A. A. Erumban, B. Los, R. Stehrer, and G. J. de Vries, 2014, “Slicing Up Global Value Chains,” *Journal of Economic Perspectives*, 28, 99–118.

Tintelnot, F., K. Kikkawa, M. Mogstad, and E. Dhyne, 2017, “Trade and Domestic Production Networks,” unpublished manuscript, available at http://felix-tintelnot.wdfiles.com/local-files/research/TKMD_draft.pdf.

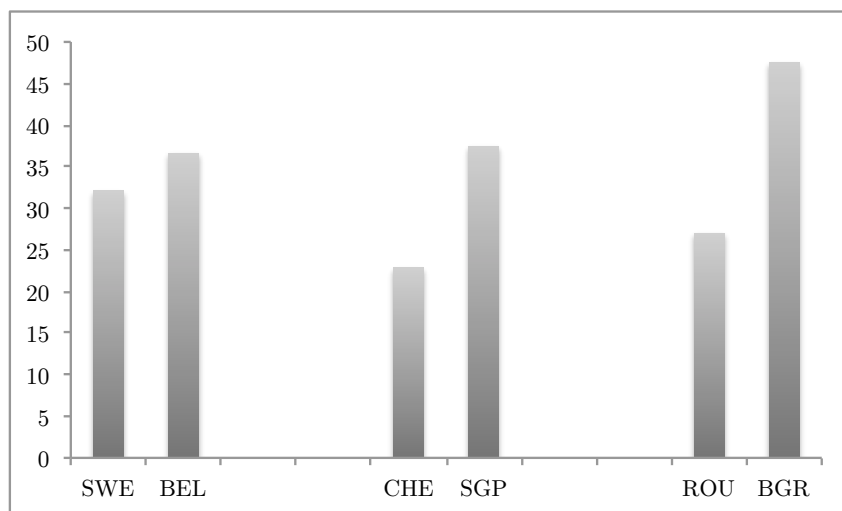
Voigtländer, N., 2014, “Skill bias magnified: Intersectoral linkages and white-collar labor demand in US manufacturing,” *Review of Economics and Statistics*, 96, 495–513.

Wang, Z., S.-J. Wei, X. Yu, and K. Zhu, 2017, “Characterizing Global Value Chains,” NBER Working Paper No. 23261, National Bureau of Economic Research.

Wang, Z., S.-J. Wei, and K. Zhu, 2013, “Quantifying International Production Sharing at the Bilateral and Sector Levels,” NBER Working Paper No. 19677, National Bureau of Economic Research. Revised March 2014.

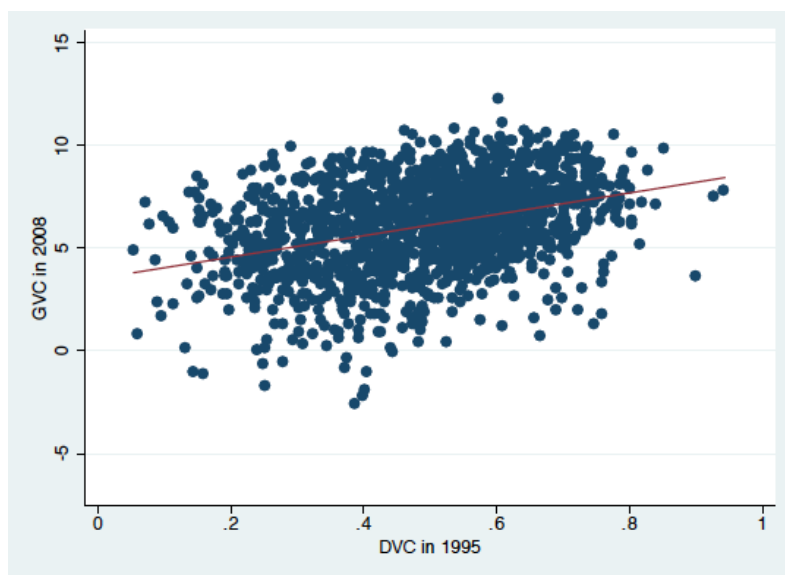
Figures

Figure 1: GVC integration of matched countries



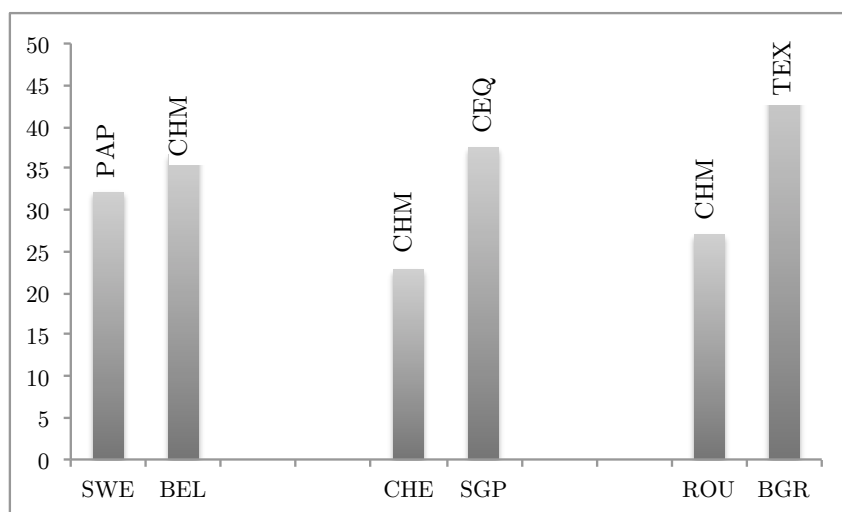
Notes: OECD data for 2008. Data are averaged across industries. GVC integration is measured as foreign value added in exports (normalized by total exports). Country names are displayed in Table A-1.

Figure 2: Correlation scatterplot between GVC in 2008 and DVC in 1995



Notes: OECD data. Each point is a country-industry combination. The GVC measure is from equation (2.3). The DVC measure is from equation (2.5).

Figure 3: Differences in GVC integration of matched countries – revisited



Notes: OECD data for 2008. Data are averaged across industries. GVC integration is measured as foreign value added in exports (normalized by total exports). Country names are displayed in Table A-1.

Tables

Table 1: DVC integration by country in 1995

| Top 10 | | Bottom 10 | |
|----------------|--------|------------|--------|
| Country | DVC | Country | DVC |
| China | 57.73% | Austria | 38.68% |
| Singapore | 56.03% | Israel | 38.11% |
| Czech Republic | 53.59% | Canada | 37.92% |
| Slovakia | 53.14% | Greece | 37.56% |
| Bulgaria | 51.37% | Hong Kong | 36.85% |
| Korea | 50.14% | Malta | 35.92% |
| New Zealand | 49.91% | Mexico | 34.69% |
| Estonia | 48.52% | Luxembourg | 29.76% |
| Hungary | 48.07% | Cyprus | 28.84% |
| Croatia | 47.95% | Cambodia | 23.50% |

Notes: OECD data for 1995. Data are averaged across industries. The DVC measure is from equation (2.5).

Table 2: Industries by category

| <i>LFLS</i> | <i>HFHS</i> | <i>LFHS</i> |
|----------------------------|---------------------------------|-----------------------------|
| Agriculture | Textile and leather products | Mining and quarrying |
| Food and beverages | Wood products | Pulp and paper products |
| Coke and refined petroleum | Non metallic mineral products | Chemicals |
| Basic metals | Fabricated metals | Rubber and plastics |
| | Machinery n.e.c. | Electrical machinery n.e.c. |
| | Electronic and optical products | |
| | Motor vehicles | |
| | Other transport equipment | |
| | Manufacturing n.e.c. | |

Notes: Based on the Rauch and Nunn classifications. L - low, H - high, F - fragmentation costs, S - switching costs. Industries are defined using the ISIC Rev. 3, 2 digit classification (see Table A-2).

Table 3: The net effect of DVC integration on GVC integration

| | (1) | (2) | (3) | (4) |
|--------------|-------------------------|---------------------|---------------------|---------------------|
| | Primary & Manufacturing | | All sectors | |
| <i>dvc</i> | 2.566*** [0.451] | 2.275*** [0.484] | 2.900*** [0.379] | 2.422*** [0.373] |
| <i>ca</i> | 1.519*** [0.074] | 1.099*** [0.073] | 1.472*** [0.065] | 1.025*** [0.062] |
| <i>ctau</i> | -0.002 [0.013] | -0.014 [0.019] | | |
| Observations | 1,025 | 1,025 | 1,449 | 1,449 |
| R-squared | 0.705 | 0.877 | 0.704 | 0.872 |

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Dependent variable: *fvax* (see equation (2.3)). Columns (1) and (3) include industry fixed effects and log of GDP, log of per capita GDP, and log of hub distance as controls. Columns (2) and (4) include industry and broad industry-country fixed effects. All sectors include primary, manufacturing and services.

Table 4: Industrial heterogeneity in the effect of DVC integration on GVC integration

| | (1) | (2) | (3) | (4) |
|---------------------------------|-------------------------|---------------------|---------------------|---------------------|
| | Primary & Manufacturing | | All sectors | |
| <i>dvc</i> | 2.275*** [0.484] | | 2.422*** [0.373] | |
| <i>dvc</i> \times <i>LFLS</i> | | 3.068*** [0.706] | | 3.081*** [0.690] |
| <i>dvc</i> \times <i>HFHS</i> | | 2.336*** [0.788] | | 2.552*** [0.484] |
| <i>dvc</i> \times <i>LFHS</i> | | 0.196 [0.845] | | 0.123 [0.818] |
| <i>ca</i> | 1.099*** [0.073] | 1.085*** [0.074] | 1.025*** [0.062] | 1.014*** [0.062] |
| <i>ctau</i> | -0.014 [0.019] | -0.010 [0.018] | | |
| Observations | 1,025 | 1,025 | 1,449 | 1,449 |
| R-squared | 0.877 | 0.879 | 0.872 | 0.873 |

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Dependent variable: *fvax* (see equation (2.3)). Columns (1) and (3) report the benchmark estimations of columns (2) and (4) of Table 3. All columns include industry and broad industry-country fixed effects. All sectors include primary, manufacturing and services.

Table 5: The net effect of DVC integration on GVC integration – Robustness

| Dep. Variable: | (1) Benchmark <i>fvax</i> | (2) High income <i>fvax</i> | (3) Panel <i>fvax</i> | (4) IV <i>i2p</i> | (5) GVC measure <i>fvax</i> | (6) All countries <i>fvax</i> | (7) All sectors <i>fvax</i> | (8) 2005 <i>fvax</i> | (9) 2011 <i>fvax</i> |
|------------------------------------------|---------------------------------|-----------------------------------|-----------------------------|-------------------------|-----------------------------------|-------------------------------------|-----------------------------------|----------------------------|----------------------------|
| <i>doc</i> | 2.275*** [0.484] | 1.892*** [0.716] | 0.568** [0.222] | 2.744*** [0.630] | 2.008*** [0.449] | 1.717*** [0.506] | 2.598*** [0.346] | 2.127*** [0.462] | 2.448*** [0.530] |
| <i>doc</i> \times <i>high income</i> | | 0.770 [0.867] | | | | | | | |
| <i>ca</i> | 1.099*** [0.073] | 1.098*** [0.072] | 0.240*** [0.032] | 1.103*** [0.063] | 0.749*** [0.060] | 1.167*** [0.081] | 1.090*** [0.063] | 1.210*** [0.071] | 1.157*** [0.085] |
| <i>ctau</i> | -0.014 [0.019] | -0.014 [0.019] | 0.002 [0.006] | -0.015 [0.016] | 0.020 [0.016] | 0.009 [0.024] | 0.001 [0.016] | 0.001 [0.016] | -0.014 [0.019] |
| Cragg-Donald Wald F statistic | 176.6 | | | | | | | | |
| First stage (dep. variable: <i>doc</i>) | -19.173*** | | | | | | | | |
| <i>doc</i> ^{IV} | [3.761] | | | | | | | | |
| Observations | 1,025 | 1,025 | 2,896 | 1,025 | 1,026 | 1,078 | 1,883 | 1,025 | 989 |
| R-squared | 0.877 | 0.877 | 0.748 | 0.877 | 0.884 | 0.881 | 0.858 | 0.881 | 0.873 |

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Dependent variable: *fvax* (see equation (2.3)) in columns (1)-(3) and (5)-(9); *i2p* (see equation (2.4)) in column (4). All columns excluding column (3) include industry and broad industry-country fixed effects. Column (3) includes country-year, industry-year, and industry-country fixed effects. All columns excluding column (7) exclude services industries. Number of observations in first stage regression: 1,025. *doc*^{IV} defined in footnote 26.

Table 6: Industrial heterogeneity in the effect of DVC integration on GVC integration – Robustness

| | (1) Benchmark | (2) IV | (3) Cutoff | (4) Complexity | (5) Capital intensity |
|-------------------------------|---------------------|-----------------------|-----------------------|---------------------|--------------------------|
| $dvc \times LFLS$ | 3.068*** [0.706] | 2.271*** [0.637] | 3.106*** [0.766] | 3.050*** [0.841] | 3.093*** [0.776] |
| $dvc \times HFHS$ | 2.336*** [0.788] | 4.603*** [1.025] | 1.962*** [0.694] | 2.472*** [0.684] | 1.929*** [0.694] |
| $dvc \times LFHS$ | 0.196 [0.845] | 1.169 [0.863] | 0.701 [0.930] | 1.236 [0.976] | 1.621 [1.082] |
| ca | 1.085*** [0.074] | 1.074*** [0.064] | 1.093*** [0.075] | 1.100*** [0.075] | 1.100*** [0.074] |
| $ctau$ | -0.010 [0.018] | -0.012 [0.016] | -0.012 [0.018] | -0.020 [0.020] | -0.014 [0.019] |
| Cragg-Donald Wald F statistic | | | | | |
| First stage dep. variable: | | (2.b) | (2.c) | | |
| $dvc^{IV} \times LFLS$ | | 54.65 | $dvc \times HFHS$ | $dvc \times LFHS$ | |
| $dvc^{IV} \times HFHS$ | | 0.026 [0.710] | 1.065 [0.655] | | |
| $dvc^{IV} \times LFHS$ | | -25.414*** [4.864] | 0.186 [0.769] | | |
| | | 1.533* [0.917] | -22.595*** [3.089] | | |
| | | 3.607*** [1.308] | | | |
| Observations | 1,025 | 1,025 | 1,025 | 911 | 1,025 |
| R-squared | 0.879 | 0.878 | 0.878 | 0.879 | 0.878 |

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Dependent variable: $foax$ (see equation (2.3)). All columns include industry and broad industry-country fixed effects. All columns exclude services industries. Number of observations in first stage regressions: 1,025. dvc^{IV} defined in footnote 26.

Appendices

A Data

Table A-1: OECD ICIO country coverage

| ISO3 | Country | ISO3 | Country |
|------------|------------------|------------|-----------------------|
| AUS | Australia | ITA | Italy |
| ARG | Argentina | JPN | Japan |
| AUT | Austria | KHM | Cambodia |
| BEL | Belgium | KOR | Korea |
| BGR | Bulgaria | LTU | Lithuania |
| BRA | Brazil | LUX | Luxembourg |
| BRN | Brunei | LVA | Latvia |
| CAN | Canada | MEX | Mexico |
| CHE | Switzerland | MLT | Malta |
| CHL | Chile | MYS | Malaysia |
| CHN | China | NLD | Netherlands |
| COL | Colombia | NOR | Norway |
| CRI | Costa Rica | NZL | New Zealand |
| CYP | Cyprus | PHL | Philippines |
| CZE | Czech Republic | POL | Poland |
| DEU | Germany | PRT | Portugal |
| DNK | Denmark | ROU | Romania |
| ESP | Spain | RUS | Russia |
| EST | Estonia | SAU | Saudi Arabia |
| FIN | Finland | SGP | Singapore |
| FRA | France | SVK | Slovak Republic |
| GBR | United Kingdom | SVN | Slovenia |
| GRC | Greece | SWE | Sweden |
| HKG | Hong Kong, China | THA | Thailand |
| HRV | Croatia | TUN | Tunisia |
| HUN | Hungary | TUR | Turkey |
| IDN | Indonesia | TWN | Chinese Taipei |
| IND | India | USA | United States |
| IRL | Ireland | VNM | Viet Nam |
| ISL | Iceland | ZAF | South Africa |
| ISR | Israel | | |

Note: Countries in **bold** excluded in benchmark estimations (Tables 3 and 4).

Table A-2: OECD ICIO industry coverage

| ISIC Rev. 3 | Industry code | Industry description | Broad industry |
|----------------|------------------|------------------------------------------------------|-------------------------|
| 01T05 | AGR | Agriculture | Primary |
| 10T14 | MIN | Mining and quarrying | Primary |
| 15T16 | FOD | Food products, beverages, and tobacco | Light manufacturing |
| 17T19 | TEX | Textiles, leather and footwear | Light manufacturing |
| 20 | WOD | Wood and products of wood and cork | Light manufacturing |
| 21T22 | PAP | Pulp, paper, paper products, printing and publishing | Light manufacturing |
| 23 | PET | Coke, refined petroleum products and nuclear fuel | Raw material processing |
| 24 | CHM | Chemicals and chemical products | Raw material processing |
| 25 | RBP | Rubber and plastics products | Raw material processing |
| 26 | NMM | Other non-metallic mineral products | Raw material processing |
| 27 | MET | Basic metals | Heavy manufacturing |
| 28 | FBM | Fabricated metal products | Heavy manufacturing |
| 29 | MEQ | Machinery and equipment n.e.c | Heavy manufacturing |
| 30,32,33 | CEQ | Computer, electronic and optical products | Electronics |
| 31 | ELQ | Electrical machinery and apparatus n.e.c | Electronics |
| 34 | MTR | Motor vehicles, trailers and semi-trailers | Heavy manufacturing |
| 35 | TRQ | Other transport equipment | Heavy manufacturing |
| 36T37 | OTM | Manufacturing n.e.c; recycling | Light manufacturing |
| 40T41 | EGW | Electricity, gas and water supply | Services |
| 45 | CON | Construction | Services |
| 50T52 | WRT | Wholesale and retail trade | Services |
| 55 | HTR | Hotels and restaurants | Services |
| 60T63 | TRN | Transport and storage | Services |
| 64 | PTL | Post and telecommunications | Services |
| 65T67 | FIN | Finance and insurance | Services |
| 70 | REA | Real estate activities | Services |
| 71 | RMQ | Renting of machinery and equipment | Services |
| 72 | ITS | Computer and related activities | Services |
| 73T74 | BZS | Research and development and other business services | Services |
| 75 | GOV | Public administration and defence | Services |
| 80 | EDU | Education | Services |
| 85 | HTH | Health and social work | Services |
| 90T93 | OTS | Other community, social and personal services | Services |
| 95 | PVH | Private households with employed persons | Services |

Notes: Industries in **bold** are classified as non-tradeable. Benchmark estimations (Tables 3 and 4) only include industries up to ISIC 37.

Table A-3: DVC integration, country average across industries

| ISO3 | DVC | ISO3 | DVC |
|-------------|------------|-------------|------------|
| CHN | 0.58 | CHL | 0.44 |
| SGP | 0.56 | TUR | 0.43 |
| CZE | 0.54 | SVN | 0.43 |
| SVK | 0.53 | RUS | 0.43 |
| BGR | 0.51 | IDN | 0.43 |
| KOR | 0.50 | USA | 0.43 |
| NZL | 0.50 | FRA | 0.42 |
| EST | 0.49 | PHL | 0.42 |
| TWN | 0.48 | DEU | 0.42 |
| HUN | 0.48 | CHE | 0.41 |
| HRV | 0.48 | ISL | 0.41 |
| ROU | 0.48 | MYS | 0.41 |
| POL | 0.48 | CRI | 0.40 |
| AUS | 0.47 | TUN | 0.39 |
| VNM | 0.47 | DNK | 0.39 |
| BEL | 0.47 | COL | 0.39 |
| IRL | 0.47 | NOR | 0.39 |
| ESP | 0.47 | ARG | 0.39 |
| PRT | 0.46 | AUT | 0.39 |
| ITA | 0.46 | ISR | 0.38 |
| LVA | 0.46 | CAN | 0.38 |
| FIN | 0.45 | GRC | 0.38 |
| IND | 0.45 | HKG | 0.37 |
| LTU | 0.45 | MLT | 0.36 |
| BRA | 0.44 | MEX | 0.35 |
| JPN | 0.44 | LUX | 0.30 |
| ZAF | 0.44 | CYP | 0.29 |
| GBR | 0.44 | SAU | 0.28 |
| NLD | 0.44 | BRN | 0.27 |
| THA | 0.44 | KHM | 0.24 |
| SWE | 0.44 | | |

Notes: Countries are ranked by decreasing DVC integration. Only countries included in benchmark estimations (Tables 3 and 4) are reported. The DVC measure is defined in equation (2.5). Country names are displayed in Table A-1.

Table A-4: DVC integration, industry average across countries

| ISIC Rev. 3 | Industry code | DVC |
|-------------|---------------|------|
| 34 | MTR | 0.70 |
| 15T16 | FOD | 0.69 |
| 27 | MET | 0.66 |
| 17T19 | TEX | 0.64 |
| 20 | WOD | 0.63 |
| 25 | RBP | 0.62 |
| 35 | TRQ | 0.61 |
| 24 | CHM | 0.60 |
| 36T37 | OTM | 0.60 |
| 30,32,33 | CEQ | 0.59 |
| 29 | MEQ | 0.59 |
| 31 | ELQ | 0.58 |
| 21T22 | PAP | 0.58 |
| 28 | FBM | 0.54 |
| 26 | NMM | 0.52 |
| 23 | PET | 0.49 |
| 60T63 | TRN | 0.44 |
| 01T05 | AGR | 0.41 |
| 65T67 | FIN | 0.38 |
| 73T74 | BZS | 0.36 |
| 50T52 | WRT | 0.35 |
| 72 | ITS | 0.35 |
| 10T14 | MIN | 0.32 |
| 71 | RMQ | 0.31 |
| 64 | PTL | 0.31 |

Notes: Industries are ranked by decreasing DVC integration. Only industries included in benchmark estimations (Tables 3 and 4) are reported. The DVC measure is defined in equation (2.5). Industry descriptions are displayed in Table A-2.

Table A-5: Summary statistics for main variables

| Variable | Observations | Mean | Std. Dev. | Min | Max |
|-------------|--------------|------|-----------|-------|-------|
| <i>dvc</i> | 1025 | 0.54 | 0.13 | 0.06 | 0.94 |
| <i>fvax</i> | 1025 | 6.28 | 2.13 | -2.65 | 12.23 |
| <i>ca</i> | 1025 | 0.36 | 0.48 | 0.00 | 1.00 |
| <i>ctau</i> | 1025 | 4.92 | 4.48 | 0.00 | 30.54 |

Notes: In-sample statistics based on benchmark estimations (column (2) of Table 3). The variable *dvc* (see equation (2.5)) is for the year 1995. The variable *fvax* (see equation (2.3)) is for the year 2008.

B Construction of cost proxies

B.1 Fragmentation costs

Rauch (1999)'s product differentiation

As discussed in Section 2, we use the Rauch classification as benchmark measure for fragmentation costs. The classification is available in two variants (liberal and conservative version) at the 4 digit SITC Rev. 2 level. We choose the liberal version, which we concord to ISIC Rev. 3 using a crosswalk via SITC Rev. 3 based on concordance tables provided by Eurostat.

For the analysis, this creates two problems. First, the classification is only available for the primary and manufacturing sectors. Second, due to the higher disaggregation several SITC categories are matched with each individual ISIC industry. We take the first issue into account by excluding industries in the services sector from the benchmark estimates and, in addition, by including services in robustness checks in which we label them as differentiated.

The second point is only problematic when the SITC categories concorded to ISIC industries differ in their categorization. To determine the ISIC industry's category in this case, we use global trade data from WITS at the SITC 4 digit level. We classify ISIC industries as differentiated if the trade share of differentiated SITC industries matched with them exceeds the mean value across all ISIC industries. Alternatively, we use the median value for robustness checks.

Costinot (2009)'s complexity

Costinot's measure is available for 41 unique observations covering 85 3 digit manufacturing SIC

Table B-1: Industry complexity

| ISIC Rev. 3 | Industry code | Complexity |
|-------------|---------------|------------|
| 35 | TRQ | 28.23 |
| 24 | CHM | 26.32 |
| 30,32,33 | CEQ | 25.36 |
| 29 | MEQ | 21.98 |
| 34 | MTR | 21.77 |
| 28 | FBM | 19.42 |
| 27 | MET | 18.29 |
| 21T22 | PAP | 17.29 |
| 31 | ELQ | 15.84 |
| 26 | NMM | 15.59 |
| 20 | WOD | 12.47 |
| 15T16 | FOD | 12.38 |
| 36T37 | OTM | 11.41 |
| 25 | RBP | 10.10 |
| 17T19 | TEX | 5.04 |

Notes: Industries are ranked by decreasing value of complexity. Complexity is the average number of training days a new workers requires to be prepared for a job in a given industry. The data are taken from Costinot (2009) and concorded to ISIC Rev. 3. industry descriptions displayed in Table A-2.

categories. This means several SIC industries have identical complexity values. We match these with our ISIC industries using a crosswalk via the 1987 SIC revision using concordance tables by the NBER and Statistics Canada. In cases where there are several complexity observations matched with one ISIC industry, we use the simple mean. Table B-1 presents the resulting industry level values of the complexity variable.

B.2 Switching costs

Nunn (2007)'s contract intensity

Nunn's contract intensity z is given by the share of differentiated intermediates in total intermediates that an industry sources as follows:

$$z_i = \sum_j \theta_{ji} Dif_j,$$

where $\theta_{ji} \equiv m_{ji} / \sum_j m_{ji}$, and Dif is a dummy equal to one if industry j is differentiated and m 's are generic elements of an ICIO matrix. Thus, the measure is built based on the Rauch classification and we can apply the values calculated using the methodology described above.

To simplify the empirical analysis we then transform this continuous variable into a dummy by defining each industry j as having high switching costs if z_i is at least equal to the median z of non-differentiated industries. For robustness exercises we replace this measure with capital intensity data readily available from the OECD.