NBER WORKING PAPER SERIES

GLOBAL VALUE CHAINS

Pol Antràs
Davin Chor

Working Paper 28549
http://www.nber.org/papers/w28549

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2021

An abridged version of this paper is to be published as a chapter in the 5th Edition of the Handbook of International Economics. All errors are our own. We thank Emily Blanchard, Alessandro Borin, Lorenzo Caliendo, Rob Feenstra, and Michele Mancini for enlightening email exchanges. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Global Value Chains
Pol Antràs and Davin Chor
NBER Working Paper No. 28549
March 2021
JEL No. F1,F2,F4,F6

ABSTRACT

This paper surveys the recent body of work in economics on the importance of global value chains (GVCs) in shaping international trade flows and multinational activity. On the empirical front, we begin reviewing several variants of the "macro approach" to measuring the relevance of global production sharing in the world economy, and we also offer a critical evaluation of the country- and industry-level datasets (or World Input Output Tables) that have been used to date. We next discuss the advantages and disadvantages of a burgeoning alternative "micro approach" that has instead employed firm-level datasets to document the ways in which firms have sliced up their value chains across countries. On the theoretical front, we propose an analogous dissection of the literature. First, we review a vast body of work developing country- and industry-level quantitative frameworks that are easily calibrated with World Input Output Tables, and that open the door for counterfactual exercises with minimal demands on estimation. Second, we overview micro-level frameworks that have treated firms rather than countries or industries as the relevant unit of analysis, and that have unveiled a number of distinctive mechanisms by which GVCs shape the determinants and consequences of international trade flows in ways distinct from traditional models of international trade. We close this survey with a discussion of a still infant literature on the desirability and effects of trade policy in a world of GVCs.

Pol Antràs
Department of Economics
Harvard University
1805 Cambridge Street
Littauer Center 207
Cambridge, MA 02138
and NBER
pantras@fas.harvard.edu

Davin Chor
Tuck School of Business
Dartmouth College
100 Tuck Mall
Hanover, NH 03755
and NBER
davin.chor@dartmouth.edu

Contents

1 Introduction

Since the early 1980s, the world economy has witnessed a significant transformation in the structure of international trade flows, giving rise to what some have called – in a somewhat Hobsbawm-like manner – the "Age of Global Value Chains" (Amador and Di Mauro, 2015; World Bank, 2020). This transformation was fueled by the combination of the information and communication technology (ICT) revolution, an acceleration in the rate of reduction in man-made trade barriers (via dozens of preferential trade agreements and China's accession to the WTO in 2001), and by political developments that brought about a remarkable increase in the share of world population participating in the capitalist system (Antràs, 2016). These forces worked in tandem to increase the extent to which firms used foreign parts and components in their production processes, as well the extent to which intermediate input producers sold their output internationally rather than to only domestic end users. In fact, it has been estimated that trade in intermediate inputs constitutes as much as two-thirds of world trade (Johnson and Noguera, 2012). As a result, the typical "Made in" labels in consumer goods no longer do justice to the amalgam of nationalities that are represented in the value added embodied in these products.

This chapter will provide an overview of how the rise of global value chains (GVCs) has shaped and may continue to shape research in the field of international trade. To be clear, some aspects of this new wave of globalization are not particularly novel. For instance, trade in intermediate inputs, an important building block of GVCs, has been studied for several decades. In fact, a few previous chapters in the *Handbook of International Economics* have devoted specific subsections to trade in intermediate inputs. Nevertheless, with one notable exception to be discussed below, the focus in those Handbook chapters was to briefly study the robustness of standard results to the inclusion of tradeable intermediate inputs, rather than to elaborate on the distinctive predictions that arise when traded intermediate inputs are modelled.[1]

Our aim is instead to focus our attention on conceptual environments in which GVCs and intermediate input flows are salient, and overview measurement techniques, empirical approaches, and theoretical frameworks specifically tailored to these environments. Our starting point will be a broad definition of GVCs that associates a global value chain with "a series of stages involved in producing a product or service that is sold to consumers, with each stage adding value, and with at least two stages being produced in different countries" (see Antràs, 2020). According to this broad definition, a firm *participates* in a GVC if it produces at least one stage in a GVC.

This broad definition is agnostic about the specific form in which foreign value added is embodied in production – e.g., raw materials, semi-processed inputs or 'tasks' – and is also consistent with various configurations of GVCs, including simple 'spider-like' structures – in which multiple parts

---

[1] Section 3.1 of Jones and Neary (1984) discusses the complexities that arise in neoclassical trade theory when the commodity space is expanded to include intermediate inputs. That section also discusses the pioneering work of Sanyal and Jones (1982) and Dixit and Grossman (1982). Section 1.4 of Krugman (1995) studies the implications of traded and nontraded intermediate inputs for factor price equalization in models with increasing returns to scale. More recently, Costinot and Rodríguez-Clare (2014) provide a quantitative evaluation of the real income gains from trade in the presence of tradeable intermediate inputs in Section 3.4 of their survey.

and components converge to an assembly plant that exports – and 'snake-like' structures – in which value is created sequentially in a series of stages, and in which production processes cross borders multiple times often involving more than two countries (see Baldwin and Venables, 2013). Similarly, when adopting a "micro" or firm-level perspective, this broad approach does not take a stance on whether trade transactions are initiated (via sunk investments) by exporters (as in Melitz, 2003), by importers, or by both. It is also consistent with complex GVCs designed and controlled by large 'lead firms' and with more decentralized GVCs in which actors do not make production decisions pertaining to stages other than those in which they directly participate. Finally, this broad approach also encompasses a narrower definition of GVCs that emphasizes additional distinctive characteristics of the rise of GVCs, namely that GVCs often entail the exchange of highly customized inputs on a repeated basis, with the contracts governing these relationships being highly incomplete and hard to enforce, features which often lead to non-trivial firm-boundary decisions (see World Bank, 2020; Antràs, 2020).

A recurring theme of this chapter will be to study the extent to which the different potential formalizations of the emergence of GVCs are material for the determination and consequences of international trade flows. We will structure the chapter into five core sections. Sections 2 and 3 will review empirical work, Sections 4 and 5 will discuss theoretical contributions, and Section 6 will overview recent work on trade policy in a world of GVCs. Within both the empirical and theoretical blocks of the chapter, we will distinguish between "macro" approaches and "micro" approaches. In what we refer to as "macro" approaches, the unit of analysis is a country or a country-industry, and the emphasis is on understanding the quantitative importance of GVCs both in determining international trade flows, but also in shaping the implications of trade policy shocks for aggregate income and for other macroeconomic variables. On the empirical front, this "macro" approach will be associated with the construction and manipulation of World Input-Output Tables to shed light on value-added trade flows across countries and the implied degree to which production process have become globalized. On the theoretical front, this "macro" approach will focus on the development of structural interpretations of these World-Input Output Tables, with the ultimate goal of constructing more reliable tools for counterfactual analysis than those that ignore the relevance of GVCs in world trade. In what we refer to as "micro" approaches, the unit of analysis will instead be the firm. Empirically, we will review a body of work that has studied GVC participation at the firm level, and that more broadly, has pushed the view that world trade flows are best understood as the aggregation of a large number of firm-level decisions related to the destinations to which firms export their products, but also the source countries from which they procure intermediate inputs, or the 'platform' countries from which they assemble goods for distant destination countries. Theoretically, the "micro" approach is largely concerned with developing tools to solve the complex problems that firms face when designing their optimal global production decisions.

We close this Introduction with a brief note on the scope of this chapter. First and foremost, this is a survey written *by* academic economists *for* academic economists. This is of special relevance for a chapter on GVCs because this has been a subject of study in many social science disciplines and because there is a burgeoning policy literature on practical aspects of the governance of GVCs. Readers eager to get an interdisciplinary overview of academic work on GVCs can consult the recent *Handbook of Global Value Chains* (Gereffi et al., 2019), while we refer readers interested in the practical policy aspects of GVCs to the recent World Development Report on *Trading for Development in the Age of Global Value Chains* (World Bank, 2020). A particularly engaging and accessible account of the rise of GVCs is provided in Baldwin (2016). Even within academic economics research, this chapter is almost exclusively focused on the implications of GVCs for the structure of international trade flows, and their real income implications. In doing so, we will not do justice to a growing literature studying the implications of GVCs for a broader set of macroeconomic phenomena related to shock propagation, synchronization of inflation, dampening of effects of exchange rate depreciation, or trade and current account imbalances (see Chapter 4 of World Bank, 2020). We also note that our survey complements and extends previous valuable overviews of theoretical and empirical work on GVCs, which include Feenstra (1998), Johnson (2018), Chor (2019), and Antràs (2020), among others.

Finally, and although we have noted above that the phenomenon of the rise of GVCs has been largely ignored in previous chapters of the *Handbook of International Economics* – the term 'global value chain(s)' is in fact not mentioned in any of the chapters in previous volumes – there is certainly some overlap between this chapter and Chapter 2 of the 4th volume of this *Handbook* (Antràs and Yeaple, 2014), which provided an overview of work on multinational activity. Indeed, Sections 5 and 7 of Antràs and Yeaple (2014) covered various aspects of the vertical expansion of multinational companies, which is an important manifestation of the rise of GVCs. To avoid duplication, we will refer readers to that chapter when quickly reviewing some work that could easily have been covered more extensively here under the umbrella of 'global value chains'.

2 Empirical Work: "Macro" Measurement

We start by surveying the growing literature that has enabled researchers to gain an empirical handle on the importance of GVCs in world trade flows. We refer to this as a line of work on "macro" measurement, since it concerns variables at the aggregate country or country-industry level. This "macro" measurement has improved because of developments on two fronts. On the conceptual front, key contributions have been made that have clarified and expanded on concepts in value added accounting. This has facilitated decompositions of the gross output and trade flows traditionally observed in the data into components that reflect input flows in GVCs. On the data front, economists have benefited from the yeoman's work that has improved, harmonized and merged national accounts statistics across ever larger sets of countries. This has made available what is variously termed inter-country or multi-region or world input-output tables, that record cross-country, cross-industry linkages in input use.

2.1 Accounting for Value Added

Consider the problem of a researcher interested in decomposing the ultimate sources of value added embodied in a good – either a finished good or a semi-finished good-in-process – whose production has traversed multiple country borders (i.e., a GVC). When this good is observed in transit, one would typically only have information on its direct country source, but not the full set of countries and industries that went into constituting the good. In the absence of direct details about the specifics of the production process, the researcher might reasonably infer this from information about input sourcing patterns observed at a more aggregate level. This is precisely the nature of the information contained in world input-output tables (WIOTs) that make it a key object of analysis for the "macro" measurement of GVCs. Multi-region or world input-output tables – that extend domestic tables to incorporate multiple geographic units – are familiar objects for scholars of input-output analysis (see Miller and Blair, 2009, Chapter 3.3), and have more recently become a tool of necessity for economists studying GVCs.

We describe several leading methods for performing such value added decompositions. Following Johnson (2018), we distinguish between decompositions of the value embodied in: (i) final goods (observed at their ultimate point of absorption into final use); and (ii) gross exports (which comprise flows of both final and semi-finished goods). Our purpose is not to present a comprehensive set of accounting identities, but rather to focus on key conceptual issues, with an eye towards a discussion of some limitations to current "macro" measurement approaches where we envision future avenues for improvement.

We conduct this discussion around the WIOT as a core empirical building block. We establish a set of notation and index conventions here in describing the structure of a WIOT, which will carry through the rest of this chapter.

Consider an economic environment in which there are $J > 1$ countries and $S > 1$ industries. The *sub*scripts i and j index countries ($1 \leq i, j \leq J$); whenever a pair of subscripts is used (e.g., to describe a trade flow variable), the left subscript will refer to the source country, while the right subscript will refer to the destination country (so ij denotes a flow from i to j). The *super*scripts r and s index industries ($1 \leq r, s \leq S$); once again, whenever a pair of superscripts is used, the left superscript will refer to the source (or selling) industry, and the right superscript to the destination (or buying) industry.

Figure 1 illustrates the structure of a WIOT. At the center of the WIOT is a $JS \times JS$ square matrix, $\mathbf{Z}$, whose typical entry Z_{ij}^{rs} is the value of inputs from industry r in country i (arrayed in the rows of the WIOT) that is purchased by industry s in country j (arrayed in the columns).[2] Moving forward, we will (somewhat inelegantly) refer to the unit of observation in either a row or column of Z_{ij}^{rs} as a "country-industry". (We use bold characters to denote vector or matrix variables, while using un-bolded characters for scalars.)

Apart from being used as an input by other country-industries, the output of industry r in country i can also be absorbed in final-use (i.e., in consumption or investment). This information is

---

[2] To be precise, Z_{ij}^{rs} is the entry in the $((i-1) \times J + r)$-th row and $((j-1) \times J + s)$-th column of the matrix $\mathbf{Z}$.

Figure 1: The Structure of World Input-Output Tables

			Input use & value added							**Final use**			**Total use**
			Country 1			$\cdots$	Country J			Country 1	$\cdots$	Country J	
			Industry 1	$\cdots$	Industry S	$\cdots$	Industry 1	$\cdots$	Industry S				
		Industry 1	Z_{11}^{11}	$\cdots$	Z_{11}^{1S}	$\cdots$	Z_{1J}^{11}	$\cdots$	Z_{1J}^{1S}	F_{11}^{1}	$\cdots$	F_{1J}^{1}	Y_1^1
Intermediate	Country 1	$\cdots$	$\cdots$	Z_{11}^{rs}	$\cdots$	$\cdots$	$\cdots$	Z_{1J}^{rs}	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
		Industry S	Z_{11}^{S1}	$\cdots$	Z_{11}^{SS}	$\cdots$	Z_{1J}^{S1}	$\cdots$	Z_{1J}^{SS}	F_{11}^{S}	$\cdots$	F_{1J}^{S}	Y_1^S
inputs	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	Z_{ij}^{rs}	$\cdots$	$\cdots$	$\cdots$	$\cdots$	F_{ij}^{s}	$\cdots$	Y_j^s
		Industry 1	Z_{J1}^{11}	$\cdots$	Z_{J1}^{1S}	$\cdots$	Z_{JJ}^{11}	$\cdots$	Z_{JJ}^{1S}	F_{J1}^{1}	$\cdots$	F_{JJ}^{1}	Y_J^1
supplied	Country J	$\cdots$	$\cdots$	Z_{J1}^{rs}	$\cdots$	$\cdots$	$\cdots$	Z_{JJ}^{rs}	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
		Industry S	Z_{J1}^{S1}	$\cdots$	Z_{J1}^{SS}	$\cdots$	Z_{JJ}^{S1}	$\cdots$	Z_{JJ}^{SS}	F_{J1}^{S}	$\cdots$	F_{JJ}^{S}	Y_J^S
Value added			VA_1^1	$\cdots$	VA_1^S	VA_j^s	VA_J^1	$\cdots$	VA_J^S				
Gross output			Y_1^1	$\cdots$	Y_1^S	Y_j^s	Y_J^1	$\cdots$	Y_J^S	$\cdots$			

reported in the set of final-use columns to the right of the $\mathbf{Z}$ matrix. We define $\mathbf{F_j}$ to be the $JS \times 1$ column vector that stacks the values F_{ij}^r of output from industry r in country i that is absorbed in country j for final-use.[3] It will be convenient to further define $\mathbf{F} = \sum_j \mathbf{F_j}$ to be the final-use vector after summing over all destination countries. For the purposes of the accounting decompositions described below, the entries of $\mathbf{Z}$ and $\mathbf{F}$ are data objects taken as given from a WIOT. Moving beyond accounting though, it should be stressed that the input and final use values reported in a WIOT should more properly be viewed as endogenous variables, as these are the outcomes of firm-level decisions over how to optimally structure input sourcing and production processes. The line of work on "macro" models of GVCs we discuss in Section 4 will emphasize this perspective.

The starting point of most value added decompositions is a basic gross output accounting identity. The gross output, Y_i^r, of a given country-industry can be expressed as the sum of the value that is: (i) absorbed in final-use; and (ii) purchased for use as an input (across all possible country-industries):

$$\begin{aligned} Y_i^r &= \sum_j F_{ij}^r + \sum_{j,s} Z_{ij}^{rs} &=& \sum_j F_{ij}^r + \sum_{j,s} a_{ij}^{rs} Y_j^s \\ & &=& \sum_j F_{ij}^r + \sum_{j,s} a_{ij}^{rs} \sum_k F_{jk}^s + \sum_{j,s} \sum_{k,t} a_{ij}^{rs} a_{jk}^{st} \sum_l F_{kl}^t + \ldots . \end{aligned} \tag{1}$$

In the second equality above, the term Z_{ij}^{rs} has been rewritten as $a_{ij}^{rs} Y_j^s$, where $a_{ij}^{rs} = Z_{ij}^{rs} / Y_j^s$ is the *direct requirements* coefficient, this being the value of the input in question (from industry r in country i) that is used in the production of \$1 of output (for industry s in country j). The final line in (1) then iteratively substitutes for Y_j^s, using the expression for gross output in each country-industry implied by the initial accounting identity. Note that the n-th term in the infinite sum in (1) is the gross output from industry r in country i that is ultimately absorbed in final-uses after exactly n production stages, under the convention that each input- or final-use transaction corresponds to a single stage. In matrix form, the above identity can be stacked across country-industries and

---

[3] More precisely, $\mathbf{F_j}$ is the column vector whose $((i-1) \times J + r)$-th entry is equal to F_{ij}^r.

expressed as:

$$\mathbf{Y} \;=\; \mathbf{F} + \mathbf{AF} + \mathbf{A}^2\mathbf{F} + \ldots, \tag{2}$$

where $\mathbf{Y}$ is the $JS \times 1$ vector of gross output values Y_i^r, and $\mathbf{A}$ is the $JS \times JS$ matrix of direct requirement coefficients.[4] From (2), one can see that $\mathbf{A}^n\mathbf{F}$ (where $n > 0$) is precisely the vector of gross output values that is absorbed in final use after traversing exactly $(n + 1)$ stages.

Although the preceding derivation may appear to be nothing more than the application of a familiar accounting identity, it is worth recognizing and emphasizing that these steps are *not* free of modeling assumptions about the nature of the underlying production technologies in this input-output system. In particular, there is a running assumption that for any given industry s in country j, the same set of direct requirements coefficients in the $((j-1) \times J + s)$-th column of $\mathbf{A}$ describes the production technology that is applied, both when that output is purchased as an intermediate input and when it is purchased for final-use. The same technology is used moreover regardless of the destination country to which that output is being sold. This latter assumption has been called into question by de Gortari (2019), a criticism that we will return to later in Section 2.2.

It is useful to introduce a complementary accounting approach that builds off an alternative gross output identity. The approach in (1) adopts a demand-driven perspective, as it is based on the sources of demand (whether as an input or for final-use) for the output of a given country-industry. An alternative approach would be to instead view gross output as the sum total of value added (i.e., payments to primary factors of production) and the intermediate inputs used in its production. Note from Figure 1 that the value added V_j^s that enters directly into the production of gross output in industry s in country j is recorded in the WIOT entries below the input-use matrix $\mathbf{Z}$. Under this supply-driven perspective (see for example Miller and Blair, 2009, Chapter 12), gross output in industry s in country j is the sum of the value of: (i) its direct payments to primary factors V_j^s; and (ii) inputs it purchases from all other country-industries:

$$\begin{aligned} Y_j^s \;=\; V_j^s + \sum_{i,r} Z_{ij}^{rs} \;&=\; V_j^s + \sum_{i,r} b_{ij}^{rs} Y_i^r \\ &=\; V_j^s + \sum_{i,r} b_{ij}^{rs} V_i^r + \sum_{i,r}\sum_{h,q} b_{hi}^{qr} b_{ij}^{rs} V_h^q + \ldots. \end{aligned} \tag{3}$$

The b_{ij}^{rs}'s that appear after the second equality sign are defined by: $b_{ij}^{rs} = Z_{ij}^{rs}/Y_i^r$; this is the share of output in industry r in country i (i.e., the *source* country-industry) that is purchased for use as an intermediate input by industry s in country j, otherwise known as the *allocation* coefficient. (This is not to be confused with the direct requirements coefficient, which instead expresses the input value as a share of the gross output of the *destination* country-industry.) The second line in (3) then performs an analogous iterative substitution in order to express gross output as an infinite

---

[4] To be more precise, the $((i-1) \times J + r)$-th entry of $\mathbf{Y}$ is equal to Y_i^r, while the entry in the $((i-1) \times J + r)$-th row and $((j-1) \times J + s)$-th column of $\mathbf{A}$ is a_{ij}^{rs}.

sum of value added terms. This can be written more compactly as:

$$\mathbf{Y} \;=\; \mathbf{V} + \mathbf{B}\mathbf{V} + \mathbf{B}^2\mathbf{V} + \ldots, \tag{4}$$

where $\mathbf{V}$ is the $JS \times 1$ vector that is the transpose of the row vector of V_j^s's, and $\mathbf{B}$ is the matrix of allocation coefficients. Note now that the term $\mathbf{B}^n\mathbf{V}$ ($n > 0$) in (4) is the value of gross output accrued from primary sources of value added that enter into production exactly $(n+1)$ stages prior.

An immediate implication of (2) is that the gross output vector can be solved for as: $\mathbf{Y} = [\mathbf{I} - \mathbf{A}]^{-1}\mathbf{F}$, where $\mathbf{I}$ is the $JS \times JS$ identity matrix. This is the familiar property that the gross output generated to yield the vector $\mathbf{F}$ for final-use can be computed by pre-multiplying $\mathbf{F}$ with the Leontief inverse matrix (Leontief, 1986). Somewhat less familiar is the fact that one can also obtain gross output in this input-output system via (4), as: $\mathbf{Y} = [\mathbf{I} - \mathbf{B}]^{-1}\mathbf{V}$. The matrix $[\mathbf{I} - \mathbf{B}]^{-1}$ is also known as the Ghosh inverse, and is the analogue of the Leontief inverse constructed instead with the matrix of allocation coefficients (Ghosh, 1958).

2.1.1 Value Added In Final Goods

When a final good from industry s is observed in the country j where it is absorbed, it embodies value added that could have originated in any (or even all) of the JS country-industries in this world economy. Drawing on tools from input-output analysis, Johnson and Noguera (2012) develop an approach for decomposing the ultimate sources from which this value added originated. To implement this, one would: (i) take the vector $\mathbf{F_j}$ of final-goods absorbed in country j; (ii) use the Leontief inverse $[\mathbf{I} - \mathbf{A}]^{-1}$ to back out the gross output needed to generate this final-use vector; and (iii) pre-multiply this by a vector of value-added shares in gross output, $\hat{\mathbf{V}}\hat{\mathbf{Y}}^{-1}$, where the 'hat' notation is used to denote a diagonal matrix that has the entries of the corresponding column vector along its main diagonal.[5] By computing:

$$\hat{\mathbf{V}}\hat{\mathbf{Y}}^{-1}[\mathbf{I} - \mathbf{A}]^{-1}\mathbf{F_j}, \tag{5}$$

this yields a vector whose $((i-1) \times J + r)$-th entry is the value added that originates from country i, industry r that is eventually absorbed in final-use in country j. Summing across all industries r and all destinations $j \neq i$, the empirical researcher can then obtain a measure of country i's value added exports, VAX_i, this being value added that originates from the country that is ultimately absorbed in the rest of the world. Johnson and Noguera (2012) propose examining how VAX_i compares against the country's gross exports (GX_i): Intuitively, this VAX_i-to-GX_i ratio reflects how involved the country is in direct exporting versus indirect exporting (i.e., through GVCs), providing an (inverse) measure of a country's engagement in these cross-border production chains.

Two remarks are in order. First, one can further disaggregate the value added share of gross output V_j^s/Y_j^s (i.e., the diagonal entries of $\hat{\mathbf{V}}\hat{\mathbf{Y}}^{-1}$) into payments that accrue to distinct factors of production. This connects value added accounting to a vast preceding literature on factor content of

---

[5] This is not to be confused with our 'hat' notation related to counterfactual analyses in Section 4.

trade accounting. On this front, Trefler and Zhu (2010) demonstrate how to implement an empirical test of the Vanek (1968) equations, suitably modified to a world with trade in intermediates, when the empirical researcher is armed with data from a WIOT.[6] More recently, Reshef and Santoni (2019) implement an accounting decomposition of factor payments in a WIOT, to explore how much of the observed fall in the labor share of GDP can be explained by the rise of GVCs.

Second, the above approach traces the value added embodied in gross output through its *forward linkages* to final uses. One could envision taking a different approach that seeks instead to keep track of sources of value added that are built into an observed amount of gross output. Toward this end, one can show that:

$$\begin{aligned}\hat{\mathbf{V}}\hat{\mathbf{Y}}^{-1}[\mathbf{I}-\mathbf{A}]^{-1}\mathbf{F_j} &= \hat{\mathbf{V}}[(\mathbf{I}-\mathbf{A})\hat{\mathbf{Y}}]^{-1}\mathbf{F_j} \\ &= \hat{\mathbf{V}}[\mathbf{I}-\hat{\mathbf{Y}}^{-1}\mathbf{A}\hat{\mathbf{Y}}]^{-1}\mathbf{Y}^{-1}\mathbf{F_j} \\ &= \hat{\mathbf{V}}[\mathbf{I}-\mathbf{B}]^{-1}\hat{\mathbf{Y}}^{-1}\mathbf{F_j}, \qquad (6)\end{aligned}$$

where we make use of the fact that the direct requirements and allocation matrices for a given input-output system are related by: $\mathbf{B}=\hat{\mathbf{Y}}^{-1}\mathbf{A}\hat{\mathbf{Y}}$. In other words, value added exports can be computed by taking the share of gross output that is absorbed by final demand in country j, $\hat{\mathbf{Y}}^{-1}\mathbf{F_j}$, and tracing its *backward linkages* to sources of value added by pre-multiplying by $\hat{\mathbf{V}}[\mathbf{I}-\mathbf{B}]^{-1}$. What the above matrix algebra steps show is that the forward and backward linkage approaches yield equivalent expressions for value added exports.

2.1.2 Value Added In Gross Exports

A closely-related task an empirical researcher could be interested in is to unpack the sources of value added that are embodied in readily available trade data (observed "as-is"), such as a country's gross exports. This is in fact the task that the early literature on the measurement of GVC activity embarked on: Hummels et al. (1998) and Hummels et al. (2001) posed this as a question of determining how much (as a share) of the dollar value of a country's gross exports can be attributed to imported intermediate inputs. This well-known measure of "vertical specialization" (VS) – or the import content in exports – was arguably one of the first measures of GVC participation, in the manner in which it captures the importance of trade flows that are involved in at least two border crossings.

Compared to the preceding exposition on the Johnson-Noguera VAX measure, efforts to decompose gross exports are by necessity more involved. This is because gross exports (GX) comprise not just shipments of final goods, but also shipments of intermediate inputs that need to be carefully tracked in the accounting. There is now an extensive line of work on gross export decompositions that can at times be quite intricate. The purpose of this subsection is not to provide

---

[6] Earlier studies on the factor content of trade, such as Davis and Weinstein (2001), acknowledged that "some error surely arises from the fact that we have assumed that intermediates used in the production of exportables are all of national origin" (p.1444), but did not tackle this issue head-on due to data limitations. See Reimer (2006) for an effort at factor content accounting with trade in intermediates in a low-dimensional setting.

an exhaustive taxonomy of these accounting identities and formulae; the interested reader is referred in particular to Koopman et al. (2014) and Borin and Mancini (2019). The discussion below instead seeks to draw out key insights from this work that a researcher would need to be alert to, in order to inform his/her choices over which measure of trade in value added or GVC participation would be most appropriate for the research question or application at hand.

Suppose that one is presented with the $S \times 1$ vector of country i's gross exports, $\mathbf{GX_i}$.[7] The schematic in Figure 2 below, adapted from Koopman et al. (2014), provides a useful organizing framework for decomposing sources of value added in these gross exports. At a high level, the dollar value recorded in $\mathbf{GX_i}$ is composed of content that is either of domestic or foreign origin. A first conceptual issue to pay attention to is that the domestic (respectively, foreign) content of exports is not equivalent to the domestic (respectively, foreign) value added embodied in gross exports. In an age of active trade of goods-in-process within GVCs, the value added that is contained in a particular input would get recorded in country i's gross exports multiple times ("double-counting") if it is shipped out of that country more than once over the course of production. For example, suppose that iron ore mined in Canada is exported to the US for fabrication into a car chassis; that chassis is then shipped back to Canada for final assembly, before the finished car is ultimately purchased by a household in the US. The Canadian value added that was part of the original iron ore will be counted twice in Canada's gross exports (the second time as part of the value of the finished car). Intuitively then, how much value added gets double-counted will depend on the extent to which the country is a hub for GVCs, with goods-in-process routed through it multiple times.[8] This breakdown of domestic (respectively, foreign) content into a pure value added component and a double-counting piece is illustrated in the second level of the schematic.

It is instructive to dive into some details of how the domestic value added (or DVA) embodied in a country's exports can be computed (less the double-counting term), by drawing on the forward linkage approach laid out in Section 2.1.1. Define $\mathbf{DVA}_i$ to be the $S \times 1$ vector whose r-th entry is the value added in country-i, industry-r exports that originates from domestic (i.e., country i) sources. $\mathbf{DVA_i}$ can then be computed as:

$$\mathbf{\hat{V}_i \hat{Y}_i^{-1} [I - A_{ii}]^{-1} GX_i}. \tag{7}$$

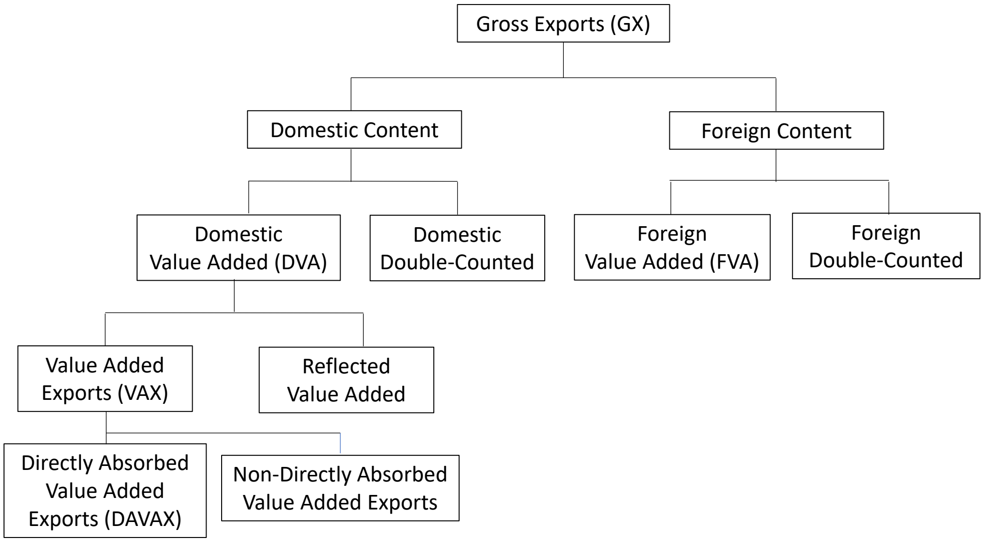
Here, $\mathbf{A_{ii}}$ is the $S \times S$ matrix block of direct requirements coefficients (along the main diagonal block

---

[7] From a WIOT, the r-th entry of $\mathbf{GX_i}$ can be computed as $\sum_{j \neq i} \sum_s Z_{ij}^{rs} + \sum_{j \neq i} F_{ij}^r$.

[8] Although double-counting could be a significant phenomenon in some specific cross-border production chains, it turns out to be relatively small in practice at the country level. In the World Input-Output Database (2013 release), when aggregating across the years 1995-2011, domestic double-counting never constitutes more than 2% of the domestic content of exports for each country; the largest shares are seen for the rest of the world aggregate (1.8%), Germany (1.5%), and China (0.9%). Similarly, the countries with the highest foreign double-counting shares in the foreign content of their exports are: Germany (1.7%), the rest of the world (1.7%), and China (1.2%). Based on authors' own calculations from the dataset made available with Chapter 1 of the World Development Report (World Bank, 2020). In Figure 3, we moreover find the share of GVC trade in gross exports, $(GX - DAVAX)/GX$, and the share of GVC trade in domestic value added, $(DVA - DAVAX)/DVA$, to be very similar, suggesting that double-counting terms that are present in GX but that are entirely removed in DVA have a relatively small effect on the level of these GVC measures, at least when computed at the world level.

Figure 2: Decomposing Sources of Value Added in Gross Exports



of $\mathbf{A}$) that corresponds to purely domestic input-output transactions (i.e., that are confined within country i's borders). The expression in (7) is reminiscent of (5): One uses the Leontief inverse of $\mathbf{A_{ii}}$ to compute the sum total of purely domestic sources of gross output – that has not crossed borders at any stage of production – that go into generating the gross export vector $\mathbf{GX_i}$. The domestic value added $\mathbf{DVA_i}$ can then be extracted from $[\mathbf{I} - \mathbf{A_{ii}}]^{-1}\mathbf{GX_i}$ by pre-multiplying it by $\mathbf{\hat{V}_i}\mathbf{\hat{Y}_i}^{-1}$, the diagonal matrix of value-added shares in each industry in country i. For the country as a whole, domestic value added in exports (DVA_i) is simply the column sum of $\mathbf{DVA_i}$. The foreign value added vector ($\mathbf{FVA_i}$) is in turn given by the residual: $\mathbf{GX_i} - \mathbf{\hat{V}_i}\mathbf{\hat{Y}_i}^{-1}[\mathbf{I} - \mathbf{A_{ii}}]^{-1}\mathbf{GX_i}$.

Two properties of domestic value added in exports make it a concept of particular interest. First, the non-domestic value added share of gross exports, $(GX_i - DVA_i)/GX_i$, is precisely equal to the vertical specialization measure – the import content in gross exports – formulated by Hummels et al. (2001). See in particular Johnson (2018) for this result in a two-country case, and Borin and Mancini (2019) for a more general proof. Second, Los et al. (2016) show that DVA_i is equal to the decrease in country i GDP that would be implied by the input-output system if exports of country i to the rest of the world – both of intermediate inputs and final goods – are shut down, holding all other entries of the WIOT constant. Such a counterfactual shift is not easy to rationalize in a fully-specified model with general equilibrium adjustments. But this "hypothetical extraction" approach nevertheless provides an intuitive interpretation for DVA_i, as well as a convenient method for calculating it that sidesteps the need to work with accounting decompositions.

Is domestic value added (DVA) in exports then also equivalent to value added exports (VAX) from Johnson and Noguera (2012)? The answer is no, as pointed out by Koopman et al. (2014). This is because DVA in general contains domestic value added that eventually gets "reflected" back and absorbed in final uses in the home country, as shown in the third layer of the schematic in Figure 2. (Using the stylized car GVC example from above, US value added embodied in its chassis exports to Canada would be reflected back in the assembled car that is purchased by the US household.) Naturally, the gap between DVA and VAX is likely to be larger for countries that are engaged in GVCs with a lot of to-and-fro trade across borders in parts and components, such as in the model of Yi (2010), and that also absorb a significant amount of the finished goods back within their domestic economies.[9]

We illustrate in Figure 2 a final layer that breaks down VAX into two components. As shown by Borin and Mancini (2019), one can explicitly keep track of the part of VAX that crosses exactly one country border ($DAVAX$, or "directly absorbed" VAX): this comprises value added that is immediately absorbed in final use in its first destination country, or that is used as an input in production processes that are contained entirely in that destination country. The difference between gross exports and $DAVAX$ is value which makes at least two border crossings, and can therefore be regarded as trade flows involved in GVCs. Building on this, Borin and Mancini (2019) proposed examining $(GX - DAVAX)/GX$ as a measure of the share of "GVC trade" in gross exports. This in turn can be written as the sum of two pieces: a first that captures forward production linkages, given by $(DVA - DAVAX)/GX$, this being domestic value added in exports that is used abroad and then re-exported (hence, crossing two borders); and a residual term $(GX - DVA)/GX$, which can be interpreted as capturing imported content in backward production linkages.[10]

Before turning to the patterns in the data, we highlight a subtle accounting issue that becomes relevant if one is seeking to decompose *not* gross country exports, but rather exports observed at other levels of disaggregation (such as bilateral or industry or country-by-industry trade flows).[11] One might imagine that this amounts to taking the decomposition of gross country exports just described, breaking it down into finer terms, and then assigning these accordingly to each destination country or industry bin. As it turns out, a consequence of double-counting is that there is not a unique way to perform this assignment. A modified version of the car GVC example will help to clarify this: Suppose that the iron ore from Canada is first exported to Mexico (rather than the US) to produce the chassis; this is then exported back to Canada for assembly, after which the car is finally sold in the US. One could either label the value added in the initial iron ore exports to Mexico as DVA, with that same content in Canada's exports of the finished car to the US then being labeled as double-counting, or vice versa. This distinction clearly does not matter if we are decomposing Canada's gross exports as a whole. However, when the object of interest is bilateral, industry or country-industry exports, the value added in question would be classified as DVA in mining exports to Mexico and as double-counting in motor vehicle exports to the US under the

---

[9] This is broadly consistent with what we see in the data: the reflected share of domestic value added is 8.6% for the US, 5.5% for the rest of the world, and 3.0% for Germany, these being the three countries with the largest reflection shares over 1995-2011 in the World Input-Output Database (2013 release). Based on authors' own calculations from the dataset made available with Chapter 1 of the World Development Report (World Bank, 2020).

[10] The latter "backward linkage" term is closely-related, though not strictly equivalent, to the Hummels et al. (2001) vertical specialization measure; see Borin and Mancini (2019) for a detailed discussion.

[11] See for example Wang et al. (2013). Los and Timmers (2018) take a different approach to this question by focusing on value added exports at the bilateral level.

former convention; under the latter approach, it would instead be treated as double-counting in mining exports to Mexico and as DVA in motor vehicle exports to the US. The upshot is that decompositions of export flows at these more detailed levels – if consistently performed – require that one specify an accounting convention. There are two natural choices here: a *source-based* approach, where the value added is labeled as DVA the first time it exits the domestic country (and is treated as double-counting thereafter); or a *sink-based* approach, where the value added is instead classified as DVA the final time it exits the country's borders (Nagengast and Stehrer, 2016; Borin and Mancini, 2019).

2.1.3 "Macro" Trends in GVC Activity

We next illustrate several broad trends in GVC activity in recent decades, with measures computed from the underlying dataset in Chapter 1 of the World Development Report (World Bank, 2020). This dataset reports gross export decompositions that have been performed on several commonly-used WIOT databases, following closely the approach summarized in Figure 2.[12] We work specifically with the decompositions for the 2013 release of the World Input-Output Database (WIOD), which provides annual observations for 35 industries and 41 countries (including a rest-of-the-world aggregate) from 1995-2011; this allows us to uncover GVC trends that were already in motion in the 1990s, that would not be reflected in the later 2016 release of the WIOD (which runs from 2000-2014). (See Section 2.2 below for a discussion of the merits and limitations of different WIOT data sources.)

We examine four measures of the prevalence of GVC activity in international trade flows. These are: (i) the vertical specialization measure VS of Hummels et al. (2001), computed as $(GX - DVA)/GX$; (ii) the ratio of value added to gross exports, VAX/GX, as proposed by Johnson and Noguera (2012); (iii) the share of GVC trade – that involved in more than one border crossing – in gross exports, $(GX - DAVAX)/GX$, from Borin and Mancini (2019); and (iv) the share of GVC trade in domestic value added in exports, $(DVA - DAVAX)/DVA$. Note that at the country level, VS is exactly equal to one minus the ratio of domestic value added in gross exports, where the latter (DVA/GX) is a measure of GVC participation that Koopman et al. (2014) have spotlighted in their work.

Measures (i), (iii) and (iv) seek in the construction of their respective numerators to capture trade flows that are involved in multiple border crossings; these three measures are thus increasing in the extent to which production is conducted within GVCs. The measure in (iv) is one we suggest, as a close counterpart to (iii), where we take the Borin-Mancini measure and remove the foreign content and domestic double-counting terms from gross exports, before assessing the importance of value added that crosses more than one border. On the other hand, the Johnson-Noguera VAX/GX ratio in (ii) is an inverse measure of GVC activity, since gross exports exceed VAX by a greater

---

[12] The data are available at: http://pubdocs.worldbank.org/en/834031570559525797/Chapter-1.zip. These have been computed at the country-by-industry-by-year level using a source-based accounting approach for double-counting terms.

extent when there is more indirect trade involving intermediate inputs; in the illustration below, we therefore subtract VAX/GX from 1, which then correlates positively with the other GVC trade measures.

Figure 3 displays the trends in these measures of GVC activity when computed for the world as a whole.[13] There are naturally differences across these measures in their average levels, given that they capture distinct aspects of GVC flows; for example, the VS measure focuses on flows involved in backward linkages, while $(GX - DAVAX)/GX$ encompasses both backward and forward linkages. What is remarkable though is the tight correlation (>0.99) across any pair of the measures, at least at this broad level of aggregation. This paints a uniform message about trends over time, regardless of one's preferred measure of GVC participation: Cross-border GVC activity rose steadily from the mid-1990s until the late 2000s. Looking across the measures, we see that the import content in exports VS increased from 0.20 in 1995 to 0.27 in 2008. Although there is a small gap between $1-(VAX/GX)$ and $VS = 1-(DVA/GX)$ due to the exclusion of reflected trade from VAX, this does not detract from the fact that both these measures display an essentially identical time trend. Similarly, the share of gross exports associated with GVC trade, $(GX - DAVAX)/GX$, rose from 0.35 in 1995 to a peak of 0.47 in 2008 (right vertical axis). When this is computed instead as the share of domestic value added that crossed multiple borders, $(DVA - DAVAX)/DVA$ as in (iv), the time pattern is preserved while the average share falls to a level comparable to the preceding VS and $1-(VAX/GX)$ measures (left vertical axis).

The above corroborates the findings of Johnson and Noguera (2017). Their estimates of the VAX/GX ratio, constructed off of input-output tables from earlier decades, indicate that this trend of rising engagement in GVCs commenced as early as the 1970s, with a sharp acceleration from 1990-2008. The global financial crisis appears to have marked a halt and even a slight reversal to this rise in GVC activity, as evident too from Figure 3, coinciding with the onset of a widely-documented slowdown in the growth of world trade as a share of GDP (see for example, the World Economic Outlook, IMF, 2016).

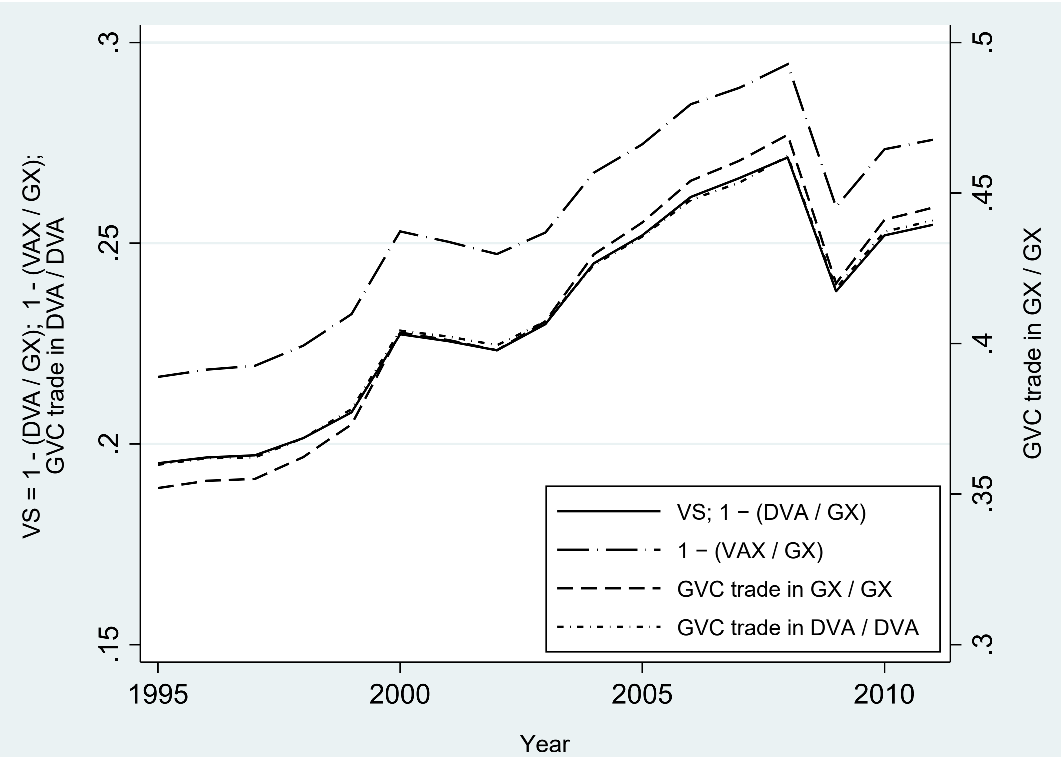
The measures at the global level mask substantial heterogeneity in the extent of participation of different countries and industries in GVC trade. We illustrate this variation with the $(GX - DAVAX)/GX$ measure.[14] When re-computed at the country level (after summing the relevant decomposition terms across all years in the sample), this Borin-Mancini measure of participation in GVC trade ranges from a low of 0.32 (Japan) to a high of Luxembourg (0.70). Not surprisingly, the observations with the highest shares of GVC trade in gross exports are small open economies deeply embedded in key regional or global supply chains: after Luxembourg, these are in descending order Slovakia (0.61), Hungary (0.60), Czech Republic (0.59), and Taiwan (0.58). On the other end of the spectrum, the countries with the lowest values of $(GX - DAVAX)/GX$ are all economies with relatively large domestic market sizes: in ascending order after Japan, these are Brazil (0.33),

---

[13] We first sum up GX, DVA, VAX, and $DAVAX$ respectively to the world level across all country-industries in a given year, and then compute the measures as defined in (i)-(iv).

[14] A similar set of takeaway messages emerges with the VS, $1-(VAX/GX)$ and $(DVA-DAVAX)/DVA$ measures (available on request).

Figure 3: GVC Trade over Time



India (0.35), the US (0.35), and Canada (0.36).[15] Looking across industries instead (over the entire 1995-2011 period), the highest values of $(GX - DAVAX)/GX$ are associated with manufacturing and mining industries, namely: Basic metals and fabricated metal (0.59), Coke, refined petroleum and nuclear fuel (0.57), and Electrical and optical equipment (0.51). On the other hand, the industries least engaged in GVCs are all services: Private households with employed persons (0.17), Social Work (0.19), Education (0.19).

Despite this rich cross-sectional variation, the pattern illustrated in Figure 3 – of rising GVC activity until 2008, and a slight turnaround thereafter – is one that pervades the time variation even at a more detailed level. In panel regressions that use the GVC participation measures constructed at the country-industry level as dependent variables, in which we include a full set of year and country-industry fixed effects, we uniformly find that the estimated year coefficients increase in magnitude up until 2008, before receding slightly; this is true for each of the measures (i)-(iv).[16]

---

[15] Interestingly, there is a fair amount of stability over time in how countries rank according to the measures of GVC participation. The country-level measures of $(GX - DAVAX)/GX$ constructed separately for 1995-1997 and 2009-2011 have a Spearman rank correlation of 0.75.

[16] To be more specific, we estimated regressions of the form:

$$GVC_{it}^{r} = \alpha_0 + \sum_{t=1996}^{2011} \alpha_t + D_i^r + \epsilon_i^t,$$

where GVC_{it}^{r} is in turn one of the GVC measures (i)-(iv) for country i, industry r, in year t; the D_i^r are country-industry fixed effects, while the α_t's are the year dummies of interest. With the Borin-Mancini measure of $(GX - DAVAX)/GX$

Why GVC activity appears to have ebbed is a question that warrants more study, particularly as newer years of WIOT data become available. An interesting question will be whether there are common underlying drivers, such as a weak recovery in global demand or rising trade frictions, that simultaneously explain both the slowdown in trade as a share of world GDP and in GVC trade as a share of exports.

We conclude this subsection by discussing several empirical applications of measures of trade in value added and GVC participation. At a basic level, this has allowed researchers and policy practitioners alike to better characterize the manner of countries' engagement in GVCs. A common distinction drawn here – which will be relevant in our discussion of models of GVCs in Section 5 – is that between backward GVC participation (i.e., the extent to which a country's exports embody value added imported from abroad) and forward GVC participation (i.e., the extent to which a country's exports are not directly absorbed in the immediate destination country, but are embodied in that country's subsequent exports). Borin and Mancini (2019) propose a particular decomposition of their GVC trade measure $(GX - DAVAX)/GX$ into terms that separately reflect these two directions of GVC participation. The World Development Report vividly illustrates the variation across countries in the extent of their backward participation in GVCs (see Map 1.1 in World Bank, 2020), while discussing how the balance between backward and forward GVC participation is a reflection of the underlying set of industries – in terms of the mix between commodities, basic manufacturing, advanced manufacturing, and innovation – that a country is currently engaged in.

Fernandes et al. (2020) undertake a wide-ranging exploration of the forces that drive this GVC participation in the data. Their findings highlight how traditional determinants of trade flows – factor endowments, institutions, geography, trade policy – correlate with backward and forward GVC participation, often over and above their role in explaining gross exports. While the patterns uncovered are informative, the extent to which one can attribute a causal interpretation is limited by the familiar challenge of identifying good instrumental variables in cross-country settings. More reassuring from the standpoint of identification, exploiting cross-country, cross-industry variation as in Romalis (2004), Fernandes et al. (2020) find that interaction terms between country factor endowments and industry factor intensities explain patterns of forward and backward participation in GVCs. This builds on earlier work by Ito et al. (2017), who point to the relevance of Heckscher-Ohlin forces for explaining the pattern of value added exports.

The 2020 World Development Report also highlights how GVC participation is associated with positive country developmental outcomes (see Chapter 3, World Bank, 2020). This includes higher growth in GDP per capita and labor productivity, gains in poverty reduction, the transfer of skills and knowhow, as well as employment creation that often benefits the female workforce. This body of evidence is admittedly more policy-oriented; while the correlations here are striking, more work remains to be done by way of investigating causal mechanisms. Altomonte et al. (2018) provides one such effort at bridging this causality gap, by proposing the use of an instrumental variable based

---

for example, we obtain: $\alpha_{1996} = 0.003$ (standard error: 0.001), $\alpha_{2008} = 0.087$ (0.005), and $\alpha_{2011} = 0.066$ (0.005), with $N = 23,887$.

on the availability of locations suitable for deep water ports that can accommodate large modern container ships; this is shown to predict well a country's DVA in exports, which in turn is positively linked with growth in income per capita.

Notably, measures of trade in value added have prompted reappraisals of several macro phenomenon. The magnitude of prominent trade imbalances, such as the U.S.' trade deficit vis-à-vis the rest of the world, and more specifically with China, is considerably smaller when assessed in value added rather than gross terms (Johnson and Noguera, 2012; Koopman et al., 2014). On a related note, efforts to understand how exchange rate movements would affect the competitiveness of a country's exports need to take into account the use of imported inputs in a GVC world. Toward this end, several authors have advocated alternative constructions of countries' real effective exchange rates (REER), that incorporate information on value added exports when weighting across individual trade partners' exchange rates (Bems and Johnson, 2017; Patel et al., 2019). Chapter 4 of World Bank (2020) provides a neat overview of this macro literature on the implications of GVCs.

2.2 Data Sources, Limitations and Directions for Improvement

The quality of the "macro" measures of GVC participation just described hinges on the reliability of the information contained in the underlying WIOT. Researchers have thus benefited from concerted efforts over the past decades to improve on the construction of WIOTs; see for example the symposium in the 2013 issue of *Economic Systems Research.* We briefly overview the methodological progress on this front, and use this as a springboard for discussing ongoing limitations and promising directions for improving on the measurement and quantification of GVC activity.

At the expense of stating the obvious, the construction of WIOTs – by such initiatives as the GTAP, OECD-ICIO, Eora, and WIOD – is an extensive undertaking. This requires at the onset bringing together domestic input-output tables across a large enough set of countries. Several practical difficulties emerge immediately. The need to harmonize across different countries' classification systems inherently limits how detailed a set of industries one can work with. Most available WIOTs thus have industries at a level of disaggregation akin to three-digit NAICS codes (in the order of about 50 industries). Moreover, national input-output tables or supply-use tables are available only with a lag, often at benchmark five-year intervals. Datasets such as the GTAP thus focus on tables in selected reference years.[17] To construct annual time series, other datasets such as the WIOD and Eora use procedures to estimate domestic flows in non-benchmark years, while respecting adding-up constraints imposed by information on industry-level aggregates – such as gross output, value added, imports, and exports – that is available from standard statistical sources at a yearly frequency. Relative to the WIOD, the Eora has made much more extensive use of computational algorithms to generate a series of tables that covers 187 countries starting in 1990, though users should note that this is accompanied by a set of reliability statistics and a documentation of instances where

---

[17] This was true too of early versions of the OECD-ICIO, although the most recent 2018 version contains annual tables for 2005-2015. Based on these inter-country input-output tables, the OECD also makes available a large set of trade in value added indicators in their TiVA dataset, including terms from the gross export decomposition illustrated in Figure 2.

conflicting constraint requirements were encountered in its construction (see Lenzen et al., 2013).

The more substantive challenge in constructing a WIOT is the task of populating its off-diagonal block entries, which relay the crucial information on cross-country flows broken down by their uses in the destination country. More specifically, how does one pin down the Z_{ij}^{rs} and F_{ij}^{r} entries for $i \neq j$, when these are not directly observed? For imports by end-uses, what is instead available is information aggregated across source countries, namely the value of country j's imports of products that map to industry r that is respectively: (i) used as an input by a given purchasing industry s, i.e., $\sum_{i \neq j} Z_{ij}^{rs}$; and (ii) absorbed in final use, i.e., $\sum_{i \neq j} F_{ij}^{r}$. The standard approach is to adopt a set of proportionality assumptions of the form:

$$\tilde{Z}_{ij}^{rs} = \omega_{ij}^{r} \sum_{i \neq j} Z_{ij}^{rs}, \text{ and } \qquad \tilde{F}_{ij}^{r} = \omega_{ij}^{r} \sum_{i \neq j} F_{ij}^{r}, \tag{8}$$

where the ω_{ij}^{r}'s are weights that reflect the importance of country i as a source of these imports. For example, Johnson and Noguera (2012) set ω_{ij}^{r} equal to the share of imports from i in country j's total industry-r imports, while Trefler and Zhu (2010) express these imports as a share of j's absorption (output less net exports) in industry r.[18]

An inherent assumption in (8) is that the breakdown of flows by country of origin is identical regardless of the intended end-uses of the imports in question. In practice though, the composition of products in say the automobile industry that are imported as an input (e.g., car parts) could well differ from that which is imported for final use (e.g., assembled cars). More recent WIOT initiatives have taken steps to relax this feature of the proportionality assumptions, by bringing in the UN Broad Economic Categories (BEC) classification system to distinguish between products by end-use categories, namely as intermediate inputs versus final (consumption or investment) goods. One can then construct from the product-level import data a separate set of ω_{ij}^{r} weights for each of these two broad end-use categories, to apportion observed flows of imported intermediates and final goods respectively.[19] This approach has been adopted in the construction of the WIOD (Dietzenbacher et al., 2013), the OECD-ICIO (Koopman et al., 2014), as well as in the most recent versions of the GTAP (Carrico et al., 2020).[20] While clearly a useful step in the right direction, note though

---

[18] In the context of the factor content of trade literature, Puzzello (2012) finds that such proportionality assumptions do not appear to generate a large bias if the purpose is to decompose the content of *net* trade. This is done by comparing the accounting results when using the Asian International Input-Output Tables (AIIO), which incorporates information from firm surveys on imported input purchases to help pin down the off-diagonal block entries, against that obtained when applying the standard proportionality assumptions.

[19] Direct information on the composition of imports of services by end-uses is even more challenging to obtain. The WIOD 2013 release uses the Eurostat import tables, computes for each service industry a simple average across countries of the share of imports by broad end-use categories, and applies this set of benchmark weights to the entire sample of countries in the WIOD. See Dietzenbacher et al. (2013) for further caveats on the quality of the WIOD data on services trade flows.

[20] This modification to the proportionality assumptions appears to result in some meaningful adjustments to measures of GVC participation. See for example Figure 2 in Timmer et al. (2015), which compares country-level VAX/GX ratios computed from different leading sources of WIOT data. While the VAX/GX measure is highly correlated across the different WIOT sources, slightly larger gaps emerge between the measures calculated from the WIOD and those from Johnson and Noguera (2012), where the latter is based on an underlying set of tables constructed using common weights across end-use categories. The gaps are more noticeable among countries with low

that a common set of ω_{ij}^{r} weights continues to be applied to imports of inputs across all purchasing industries in country j; in other words, the source-country shares of automobile industry products imported by say the Automobile manufacturing industry would be identical to that imported by the Transportation services industry.

Building on the above observations on the state-of-play on the WIOT data front, we highlight three potential directions for improvement in measurement. First, we anticipate and welcome more efforts to bring to bear more detailed micro data, to directly inform and refine the construction of the proportionality weights. The IDE-JETRO's Asian International Input-Output (AIIO) Tables is an example of work along these lines (Meng et al., 2013): Using proprietary firm-level surveys run in several Asian countries, the AIIO uses the information gathered on the composition of firm imports to construct weights that vary by the identity of the purchasing industry (i.e., ω_{ij}^{rs} weights that vary across destination industry s). Such work is likely to benefit as empirical researchers gain more access to administrative data on firm operations that can be merged with customs data on the international trade patterns of these firms. For example, efforts at value added accounting at the firm level have been executed by Kee and Tang (2016) for China and for Bems and Kikkawa (2021) for Belgium. Using a combination of Chinese customs and manufacturing survey data, Kee and Tang (2016) document a noticeable increase in the domestic value added content of firm-level exports during 2000-2007, a period of rapid trade liberalization for the country. Bems and Kikkawa (2021) work with a very rich data environment that allows them to observe not only firms' cross-border input purchases, but also domestic purchases in value added tax records. Their findings suggest that the use of sectoral aggregates in WIOTs to perform value added accounting results in an over-statement of trade in value added, as this overlooks heterogeneity in input sourcing patterns across large versus small firms.[21] Given the high demands on data, existing studies that incorporate firm-level data to improve value added accounting are either focused on individual countries, or limited in geographic coverage (in the case of the AIIO, to just ten Asia-Pacific economies). There remain significant hurdles to linking such micro datasets across countries – not least of which is how to preserve the confidentiality of firm identities when merging data across countries – though this may eventually become feasible in economically-integrated regions such as the EU where there is a history of collaboration among national statistical agencies.

A second distinct data challenge is raised by the presence of multi-product firms (Bernard et al., 2010, 2011). Large and more productive firms that tend to select into importing and exporting – and thus be a GVC participant – are also more likely to be active in manufacturing and exporting multiple products. Even if one were armed with detailed data on such firms' imported intermediates, one would still require information on how these inputs are apportioned across the manufacturing processes for different products in order to perform an accurate accounting of value added flows. Such

---

VAX/GX ratios, that are thus in principle most engaged in GVCs.

[21] A related strand of work has sought to disaggregate country input-output tables by using computational algorithms (subject to adding-up constraints), in order to distinguish between the input sourcing patterns of key subsets of firms. In the context of China, Koopman et al. (2012) use this approach to explore differences across trade regimes (in particular, processing versus ordinary trade), while Tang et al. (2020) have examined variation by firm ownership type (in particular, state- versus private-owned enterprises).

information is as of now not routinely collected in firm surveys or manufacturing censuses. While one might hypothesize that this concern could be alleviated with access to even finer establishment-level microdata – as establishments could be more specialized in their product scope – it has been well documented that multiproduct firms often feature multiproduct establishments (see for example Boehm et al., 2018, in the context of India).[22]

A third measurement issue is raised by de Gortari (2019), who documents how observed patterns of input sourcing can differ systematically across firms even within the same industry, depending on the identity of the export markets that the firms' output is destined for. Using Mexican customs data, de Gortari (2019) shows for example that motor vehicle firms in Mexico whose main export market is the US (respectively, Germany) tend to bring in a disproportionate share of their imported inputs from the US (respectively, Germany). The methodologies described thus far fail to take this phenomenon into account: When we impose a uniform set of input-output coefficients for the entire Mexican motor vehicle industry, this does not pick up differences in the sources of value added embodied in cars that are assembled for export to the US versus Germany. As a practical consequence, this would understate the extent of GVC trade between the US and Mexico in this industry, and imply lower welfare costs should a US-Mexico tariff war disrupt cross-border value chains. Addressing this shortcoming requires improvements in measurement, specifically firm-level data that would allow the researcher to construct input sourcing patterns that differ by GVCs when these are (at a minimum) distinguished by the immediate destination of firms' exports. At a deeper level though, this criticism extends beyond concerns about measurement. Conceptually, de Gortari (2019) draws attention to a key embedded assumption in standard accounting approaches, that the roundabout structure of global production can be summarized by a single technology, with a single matrix of input-output coefficients.[23] Put otherwise, value added accounting is ultimately *not* free of modeling assumptions about the structure of production within GVCs.

2.3 Measures of Positioning in GVCs

Apart from quantifying the size and share of GVC-related trade flows, researchers have taken a further interest in understanding the positioning of countries and/or industries within GVCs. We discuss here a class of such production staging measures – of "upstreamness" and "downstreamness" – that can be constructed from input-output tables. These measures provide at a descriptive level a formal basis for statements about whether a given country is specialized in relatively upstream activities, or whether its positioning is more proximate to final demand. Such notions of production staging moreover feature prominently in economic models of GVCs (see Sections 4 and 5): The positioning of countries within GVCs can be shaped by fundamentals such as productivity differences (Costinot

---

[22] Such complications related to multiproduct establishments are the reason why Boehm and Oberfield (2020) focus on single-product establishments in their study of input sourcing patterns in India.

[23] In the framework that de Gortari (2019) formulates, this translates into an assumption that the production process in GVCs is represented as a first-order Markov chain, when this should instead be viewed as a higher-order Markov chain, in which the technology coefficients can differ depending on the identity of the destination to which the output is to be sold.

et al., 2013) or geography (Antràs and de Gortari, 2020). At a more micro level, the decisions that firms make over how to structure and organize their production can differ systematically for upstream versus downstream stages within GVCs (e.g., Antràs and Chor, 2013; Alfaro et al., 2019).

We describe the formulation of these measures in the setting of a WIOT.[24] Each country-industry in a WIOT is traversed as a stage in many sequential production chains that originate in primary sources of value added and terminate when the finished goods or services are absorbed in final use. We distinguish between two production staging concepts. The first captures the positioning of a country-industry in terms of its "upstreamness" relative to sources of final demand (i.e., consumption or investment), averaged across the many production chains that connect the country-industry to final uses. The second measure on the other hand gauges the average positioning of the country-industry in terms of its "downstreamness" in relation to sources of value added (i.e., labor and other primary factors).

The measure of "upstreamness" builds upon the forward linkage decomposition of gross output presented earlier in equation (2). Recall that the n-th term, $\mathbf{A}^{n-1}\mathbf{F}$, in (2) is the vector of gross output that traverses exactly n stages to reach final demand. One can compute:

$$\mathbf{F} + 2\mathbf{AF} + 3\mathbf{A}^2\mathbf{F} + \ldots \;=\; [\mathbf{I} - \mathbf{A}]^{-2}\mathbf{F}, \tag{9}$$

from which the "upstreamness", U_i^r, of country-i, industry-r is then defined as the $((i-1)\times J+r)$-th entry of (9) divided by Y_i^r.

Intuitively, U_i^r computed in this manner is a weighted-average of the number of production stages that this country-industry's output takes to arrive at final demand. This is because the weights applied – the $((i-1)\times J+r)$-th entry of $\mathbf{A}^{n-1}\mathbf{F}$ (for $n = 1, 2, \ldots$) divided by Y_i^r – are equal to the respective shares of that gross output that traverses exactly n stages before being absorbed in final uses.[25] It is straightforward to see that $U_i^r \geq 1$, and that the minimum value of 1 is attained if and only if the entirety of the country-industry's output, Y_i^r, is absorbed directly in final demand. Moreover, U_i^r takes on a larger value when a greater share of Y_i^r is purchased as an intermediate input, and particularly so when multiple production stages are still needed before the point of final consumption/investment is reached.[26]

This upstreamness measure has several interesting properties and interpretations. Fally (2012)

---

[24] While Fally (2012) and Antràs et al. (2012) proposed the production staging measures with domestic input-output tables in mind, these extend readily to the setting of a WIOT (see Miller and Temurshoev, 2017; Antràs and Chor, 2019).

[25] In practice, one needs to account too for the value of net inventories N_i^r reported for each country-industry in a typical WIOT. The standard approach is to adopt a proportionality assumption, that the breakdown of uses of inventories across both intermediate and final use is identical to that observed in the same WIOT for non-inventorized output for each country-industry. This implies a simple correction procedure – multiplying each Z_{ij}^{rs} and F_{ij}^{rs} term in the WIOT by $Y_i^r/(Y_i^r - N_i^r)$ – before one computes the production staging measures (for details, see Antràs et al., 2012; Antràs and Chor, 2019).

[26] In terms of nomenclature, Fally (2012) refers to this measure of "upstreamness" as the "number of stages between production and final consumption". This measure of stage distance to final demand was developed contemparaneously by Antràs and Chor (2013) to test their model of firm organizational decisions along sequential production chains, though somewhat unhelpfully, they refer to it as $DownMeasure$ in that paper.

and Antràs et al. (2012) show that the U_i^r defined as above by construction from first principles is also the unique solution (up to a normalization) to the recurrence relation:

$$U_i^r = 1 + \sum_{j,s} b_{ij}^{rs} U_j^s, \tag{10}$$

where recall that $b_{ij}^{rs} \equiv Z_{ij}^{rs}/Y_i^r$ is the allocation coefficient of country-i, industry-r's use by country-j, industry-s. In other words, each country-industry is deemed according to U_i^r to be one stage more upstream than a weighted-average of all country-industries that purchase inputs from it. This provides an alternative foundation for the use of U_i^r as a measure of stage distance to final demand. Furthermore, stacking (10) across all country-industries and applying a quick matrix manipulation, one arrives at the result that U_i^r is precisely equal to the $((i-1) \times J + r)$-th row sum of the Ghosh inverse matrix, $[\mathbf{I} - \mathbf{B}]^{-1}$.[27] This establishes an equivalence between U_i^r and the measure of total forward linkages proposed by Jones (1976): Apart from capturing the average number of stages to final demand, U_i^r is also equal to the overall increase in costs that would be transmitted to total gross output in the world economy (assuming full pass-through at each stage) as a result of a unit increase in payments to primary factors in country-i, industry-r (see Chapter 12, Miller and Blair, 2009).

In an analogous fashion, the "downstreamness" from primary sources of value added is defined by working with the backward linkage gross output decomposition in equation (4). Multiplying the n-th term in (4) by n, we have:

$$\mathbf{V} + 2\mathbf{B}\mathbf{V} + 3\mathbf{B}^2\mathbf{V} + \ldots \;=\; [\mathbf{I} - \mathbf{B}]^{-2}\mathbf{V}. \tag{11}$$

The "downstreamness", D_j^s, of country-j, industry-s from primary factors is then computed by taking the $((j-1) \times J + s)$-th entry of (11) and dividing it by Y_j^s.

D_j^s is thus a weighted-average of the number of production stages traversed from primary factors to arrive at gross output in country-j, industry-s. The weights in question – the $((j-1) \times J + s)$-th term in $\mathbf{B}^{n-1}\mathbf{V}$ divided by Y_j^s – are the respective shares of gross output that accrue from value added that has passed through exactly n stages.[28] Once again, each $D_j^s \geq 1$, with equality if and only if all of the gross output of the country-industry is derived directly from primary sources of value added (with zero purchases of intermediate inputs). D_j^s is larger the greater is the use of intermediate inputs as a share of Y_j^s, and particularly so if the directly purchased inputs are themselves multiple stages removed from primary factors.

Fally (2012) and Miller and Temurshoev (2017) moreover establish that D_j^s is the unique solution

---

[27] To see this, observe that when we stack (10) across all country-industries, the recurrence relation can be written in matrix notation as: $\mathbf{U} = \mathbf{1} + \mathbf{B}\mathbf{U}$. Here, $\mathbf{U}$ is a column vector whose $((i-1) \times J + r)$-th entry is equal to U_i^r; $\mathbf{1}$ is a $JS \times 1$ vector of 1's; while $\mathbf{B}$ is the matrix of allocation coefficients. It follows that $\mathbf{U} = [\mathbf{I} - \mathbf{B}]^{-1}\mathbf{1}$, from which U_i^r is equal to the sum of entries in the $((i-1) \times J + r)$-th row of $[\mathbf{I} - \mathbf{B}]^{-1}$.

[28] Fally (2012) refers to this as the "number of production stages" embodied in the gross output of a particular country-industry.

(up to a normalization) to the recurrence relation:

$$D_j^s = 1 + \sum_{i,r} a_{ij}^{rs} D_i^r, \tag{12}$$

where $a_{ij}^{rs} \equiv Z_{ij}^{rs}/Y_j^s$ is the direct requirements coefficient of country-j, industry-s's use of inputs from country-i, industry-r. Thus, a country-industry is one stage more downstream than a weighted-average of industries that it purchases inputs from. As a consequence of (12), D_j^s can also be computed as the sum of the entries in the $((j-1) \times J + s)$-th column of the Leontief inverse matrix, $[\mathbf{I} - \mathbf{A}]^{-1}$.[29] This means that D_j^s is equivalent to the total increase in gross output in the world economy that would be required to generate the inputs to facilitate a unit increase in country-j, industry-s's output (total backward linkages).

Miller and Temurshoev (2017) and Antràs and Chor (2019) explore the empirical features of these measures of upstreamness and downstreamness, constructed using the 2013 release of the WIOD. Figure 4 below highlights a key set of stylized facts from these papers: Between 1995 and 2011, the upstreamness of country-industries rose modestly but steadily (left panel). While the median country-industry is positioned slightly more than two stages upstream from final demand, this stage distance displays a clear upward trend not just for the median observation, but also across the 25th and 75th percentiles of the distribution (within each year) of the U_i^r values. Somewhat surprisingly, the summary statistics and time trends exhibited by the downstreamness measure are very similar (right panel). The median country-industry is slightly more than two stages downstream from primary factors, though the D_j^s values display less cross-sectional dispersion than the U_i^r measure. That said, there is a distinct rise too in the downstreamness of country-industries over this sample period. Using fixed effects panel regressions, Antràs and Chor (2019) document that this upward trend in both U_i^r and D_j^s is a robust feature even in the within country-industry variation. Put otherwise, during this period when GVC activity as a whole was on the rise (cf., our earlier Figure 3), the average stage length of GVCs that each country-industry is positioned within was also increasing, as a result of an increase in both the stage distance from primary factors and the stage distance to final demand; GVCs have thus become more complex – in the sense of involving more production stages – over time.[30] Interestingly, this trend of "lengthening" GVCs tapered off after 2008, a timing that coincides with the broader slowdown described earlier in the macro measures of GVC activity.[31]

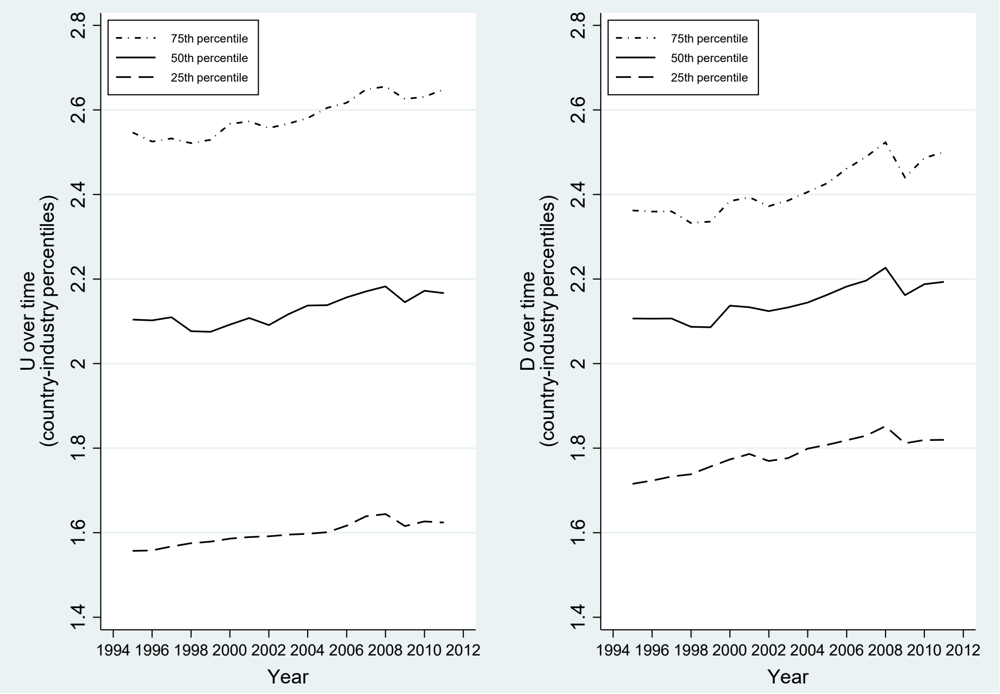
---

[29] Stacking (12) across all country-industries, the recurrence relation can be written as: $\mathbf{D}^T = \mathbf{1}^T + \mathbf{D}^T\mathbf{A}$. Here, $\mathbf{D}^T$ is the row vector whose $((j-1) \times J + s)$-th entry is equal to D_j^s; $\mathbf{1}^T$ is a $JS \times 1$ row vector of 1's; while $\mathbf{A}$ is the matrix of direct requirements coefficients. We thus have: $\mathbf{D}^T = \mathbf{1}^T[\mathbf{I} - \mathbf{A}]^{-1}$, from which D_j^s is equal to the sum of entries in the $((j-1) \times J + s)$-th column of $[\mathbf{I} - \mathbf{A}]^{-1}$.

[30] By contrast, Fally (2012) finds a decrease in the measure of downstreamness when computed for a long time series of benchmark years of the U.S. input-output tables from 1947-2002. This suggests that some within-U.S. segments of GVCs were decreasing in stage length over time, even while GVCs as a whole have been spanning more production stages since the mid-1990s.

[31] On a related note, within any given year of WIOD data, there is a strong positive correlation between upstreamness and downstreamness across country-industries (Antràs and Chor, 2019). This is driven in part by goods-producing industries exhibiting larger U_i^r and D_i^r values on average (in line with there being more production fragmentation in the manufacturing sector), while service-producing industries tend to be in chains with low U_i^r and D_i^r values (that

Figure 4: GVC Positioning (“Upstreamness” and “Downstreamness”) over Time



The above measures of production staging have found their way into a wide range of empirical applications. At a descriptive level, the measures have helped to shed light on how production and trade patterns have evolved, providing in particular a formal way of answering policy-relevant questions related to where countries are positioned within GVCs. Chor et al. (2020) use a measure of industry-level upstreamness computed from China’s input-output tables and combine these with detailed customs data, in order to infer how the global production line position of Chinese firms has shifted over time, in the absence of more direct information on these firms’ production activities within GVCs. They report a distinct rise in the upstreamness of China’s imports, at least up until 2007, even as the positioning of China’s exports became slightly more proximate to final demand; overall, this suggests that as Chinese firms have grown in size, the span of stages that they undertake within China has widened. On a related note, Li et al. (2015) find that China’s state-owned enterprises have maintained significantly higher shares of output and value added in upstream industries (such as petrochemicals and electricity generation), relative to their presence in industries that are closer to final demand.

Of note for the study of trade policy, using a broad sample of countries, Shapiro (2020) uncovers a strong negative relationship between the upstreamness of an industry and the level of protection – both tariffs and non-tariff barriers – enacted on imports from that industry. This constitutes some of the most extensive evidence to date that the pattern of applied protection is consistent with the logic of tariff escalation (see Section 6). Elsewhere, the production staging measures

---

employ primary factors and service final-users directly).

have been used to good effect to investigate the propagation of shocks across economic units. Olabisi (2020) demonstrates that industries that are more upstream display more nominal output volatility, a pattern that can be generated by a simple model of transmission of shocks to final demand. By extension, Olabisi (2020) finds too that countries whose export profiles are tilted towards more upstream industries in turn exhibit more export volatility. Separately, Liu (2019) finds that industries that are more upstream also tend to feature a higher "distortion centrality", where the latter measure captures the extent of the welfare gains that would be propagated through an input-output network when the efficiency wedges faced by the industry are reduced. With regard to the transmission of credit shocks, Kalemli-Özcan et al. (2014) find using Orbis data that firms located in more upstream sectors maintain higher levels of working capital in the form of accounts receivable, and that these levels of working capital are more sensitive to aggregate credit conditions; these patterns are consistent with the model of moral hazard in the provision of credit along production chains, developed by Kim and Shin (2012).

Looking across and within firms, a body of studies has applied these measures of production staging to test theories related to the structure and organization of GVCs. A leading though somewhat loosely-articulated theory that has drawn attention is the "smile curve" hypothesis: Based on an observation made in the early 1990s by former Acer CEO Stan Shih, this postulates that the value added in production chains is highest in the most upstream stages (e.g., research and development) and the most downstream stages (e.g., marketing and sales), with the value added contribution of midstream processing and assembly being lower in comparison. Using data from Orbis, Rungi and del Prete (2018) indeed uncover a characteristic U-shape relationship between firm-level value added and the upstreamness of their main industry of activity with respect to final demand; this complements other work that tests for a "smile curve" using broader cross-sectoral data (Ito and Baldwin, 2021).

The upstreamness measure has also been used to formally test models of firm organizational decisions along sequential production chains, such as that developed in Antràs and Chor (2013). The predictions of this model are relatively subtle (cf., Section 5.3.2): When final-good demand elasticities are sufficiently high, firms would have a greater propensity to outsource upstream stages (to better incentivize suppliers located early in the production process), before possibly integrating within firm boundaries suppliers that are positioned closer to final demand (as rent extraction motives start to dominate). A converse pattern is expected to hold when final-good demand elasticities are relatively low. Empirical tests of these predictions have been performed by using industry-level data on intrafirm import shares as a proxy for integration, and combining these with measures of the upstreamness of the imports in question (Antràs and Chor, 2013). More recent empirical work on this topic – examining the relative propensity to integrate upstream versus downstream stages – has brought in micro data, with the identity of inputs that are integrated within firm boundaries being inferred from the set of industry codes that a firm and its subsidiaries are listed to be active in (del Prete and Rungi, 2017; Alfaro et al., 2019).[32] Alfaro et al. (2019) in

---

[32] See also Bolatto et al. (2018) who examine the interplay between intangible assets and production line positioning

particular introduce a more refined measure of upstreamness in their test, that captures the stage distance of an input industry r with respect to the primary industry s of the firm headquarters (rather than measuring this with respect to final demand). This latter variable is of potential interest for empirical researchers seeking a measure of the production chain distance between pairs of industries. It is moreover equivalent to the measure of average propagation length – the average number of stages taken for a shock to be transmitted from one industry to another – formulated in Dietzenbacher et al. (2005). (See also Wang et al. (2017) and Imbs and Pauwels (2020) for several further variants and extensions of these measures of production line positioning.)

As a final and more critical remark, it should be acknowledged that the WIOT data concerns highlighted in Section 2.2 apply too to these measures of positioning within GVCs. In particular, the measures of upstreamness and downstreamness have been formulated conditional on the assumption underlying the gross output accounting identities, that there is a uniform technology matrix that fully describes the structure of roundabout production. To the extent that departures from this assumption are relevant in practice, this would suggest that there is some degree of measurement error in the upstreamness and downstreamness variables as currently constructed. Furthermore, the industry categories in most available WIOTs remain relatively coarse, which limits the usefulness of these upstreamness and downstreamness measures in terms of how much they can inform us about production-line positioning within GVCs at a more micro level. It goes without saying that a more detailed understanding of how production stages are sequenced will require that we move beyond the information aggregated in a WIOT, to work instead with firm-level data that will allow researchers to more directly observe sourcing decisions and buyer-supplier linkages.

3 Empirical Work: Micro

We turn next to survey an extensive body of work based on micro data at the firm or establishment level, that has uncovered empirical regularities and stylized facts on GVC activity as observed from the ground up. Although not all of these papers are explicitly framed as studies about GVCs *per se* – in some cases, the term "GVC" may not even appear in the paper – these have contributed to a more complete picture of the firm-level correlates of forward participation (as evidenced more broadly through exporting) and backward participation (importing) in GVCs. More recently, richer datasets have become available that track the identities of *both* parties involved in cross-border or within-country transactions. These have shed further light on the nature of buyer-supplier links within production chains and sourcing relationships.

This work uncovering micro-level evidence on GVCs is complementary to the macro measurement literature just reviewed in Section 2.1.3, with each approach having its pros and cons. The richer data environments in the micro-level studies allow for detailed investigations of mechanisms and more scope for achieving causal identification. On the other hand, virtually all existing firm-level studies are limited to data from a single country, given the practical hurdles to merging administrative

---

in global supply chains using firm-level data from Slovenia.

datasets from different countries. The view of GVCs that one obtains is thus limited to the particular segments of value chains that pass through the country in question; in particular, this makes it more difficult to infer the aggregate implications of GVC activity without making assumptions about the structure of value chain stages that operate outside the country.

3.1 Selection into GVC Participation

The advent of large firm-level datasets in research in international trade, starting in earnest in the mid-1990s, has generated a series of key stylized facts focusing on firms that export. By extension, these facts speak too to the nature of firms' forward participation in GVCs. It is now well-known that only a small share of firms engages in exporting; for example, 18% of U.S. manufacturing firms did so in 1997 (Bernard et al., 2007). These firms are nevertheless highly consequential to an economy: Exporting firms are on average larger in size and more productive than their non-exporting counterparts, and this advantage is often in place prior to the commencement of a firm's exporting episode. These patterns are consistent with the positive selection of firms along these dimensions into exporting, as more productive firms are better able to bear and amortize various fixed costs associated with attempts to enter into an export market.[33] This empirical facts about exporters have been documented extensively in developed economies (e.g., the U.S., Bernard and Jensen, 1999) as well as in developing countries (e.g. Clerides et al., 1998); see in particular Section 2 of Melitz and Redding (2014a) for an extensive review of this body of evidence.

It goes without saying that not all firms that export are necessarily engaged in forward GVC participation. One would in principle be able to more confidently label a firm as being part of a GVC if its exports were composed mainly of intermediate inputs or semi-finished goods that are being delivered to another country for further processing or assembly. That said, existing studies on firms that export typically do not distinguish between exports of final goods versus intermediates, nor are there many studies that seek explicitly to tease out the import content of firm-level exports (Kee and Tang, 2016, being a key exception).

In comparison, work on the correlates and drivers of firm-level importing has received less spotlight, even though the patterns here are no less robust and arguably no less important for understanding firms' performance. Firms that import are once again the exception rather than the rule, with only 14% of firms recorded as importers in the U.S. Census of Manufacturing in 1997 (Bernard et al., 2007). There is an analogous set of stylized facts that points to the relevance of selection into importing (and by extension into backward GVC participation), with the size and productivity advantage of importing firms over non-importers comparable in magnitude to the corresponding "premia" of exporting firms over non-exporters.[34] Importing firms that are

---

[33] This does not rule out the possibility that the converse relationship could hold too in practice, namely that firms can become more productive as a result of engaging in exporting. See for example de Loecker (2007) for evidence on "learning by exporting" in the context of Slovenia.

[34] For evidence on this, see for example Table 8 in Bernard et al. (2007) for U.S. manufacturing; Table 4 and Table 14 in Muûls and Pisu (2009) for Belgium; Table 7 in Castellani et al. (2010) for Italian manufacturing; and Appendix Table B1 in Blaum et al. (2018) for France.

larger moreover tend to purchase goods from a greater number of source countries (Antràs et al., 2017). This suggests that there are significant fixed costs associated with each source country that firms import from, which smaller firms that lack sufficient scale are not able to afford. The model-based estimates from Antràs et al. (2017) for example place these fixed costs of importing in the range of US$10,000 to US$56,000 per annum for U.S. firms depending on the source country.[35] There is moreover interesting evidence that as firms become more productive, they do not scale up their imports proportionately across all existing source countries. Blaum et al. (2019) find using French data that firms especially raise their imports from countries that are better able to produce high-quality inputs, generating what the authors term a "non-homothetic" import demand response. This builds on earlier work that has highlighted how importing firms tend to have access to more input varieties, and often end up paying higher prices – more specifically, unit values – for both domestic and imported inputs, consistent with their using higher quality inputs (see Kugler and Verhoogen, 2009, based on Colombian plant-level data).

For the study of GVCs, firms that both import and export at the same time are of particular interest. This is because these firms oversee flows of what we have termed GVC trade, namely value added that experiences more than one border crossing. In light of the above discussion on selection into importing and exporting respectively, it comes as no surprise that firms that simultaneously engage in both imports and exports are an even smaller share of all firms (11% of U.S. manufacturing firms in Bernard et al., 2007). Yet, though small and granular in terms of count, these GVC participants are clearly crucial for understanding aggregate outcomes in an economy given the disproportionate share of activity that they account for. In the U.S., firms that both import and export constituted close to a third of private-sector employment and an overwhelming share (close to 90%) of U.S. trade in 2000 (Bernard et al., 2009); these firms are moreover more likely to undertake complex sourcing strategies or to be engaged in multinational activity. Such patterns have been documented too in other countries, including in emerging economies: Using data from key Chilean manufacturing industries, Kasahara and Lapham (2013) show that the productivity distribution of firms engaging in both imports and exports is shifted to the right relative to that for firms that only import, only export, or do neither.[36]

There are strong conceptual reasons – and an accompanying body of evidence – that support the view that the respective firm-level decisions to import and export are often intrinsically linked. A growing number of studies has demonstrated that firms that raise their purchases of imported inputs tend themselves to become more productive and to expand the range of products they manufacture. This has been documented across a broad range of country settings, including Indonesia (Amiti and Konings, 2007), Chile (Kasahara and Rodrigue, 2008), India (Goldberg et al., 2010), and Hungary (Halpern et al., 2015); the latter two papers in particular point to a substantial role for the extensive

---

[35] Using a structural model of firm importing and exporting decisions with both sunk and fixed costs, Kasahara and Lapham (2013) estimate fixed costs of importing that lie in a similar ballpark – between $28,000 and $117,000 in 1990 US dollars – for Chilean firms in selected manufacturing industries. (They also recover estimates of the fixed costs of exporting between $36,000 to $83,000.)

[36] Lu (2010) can be viewed as an exception in this regard: Firms that engage in processing trade appear to exhibit a low exporter productivity premium in China.

margin of import growth, namely access to a greater set of imported input varieties, as a crucial margin behind the overall improvement in firm performance.[37] On a closely related note, Bøler et al. (2015) show using Norwegian data that there can be strong complementarities at the firm level between the decisions to import intermediate inputs and to undertake R&D, so that reductions in costs associated with importing can spur R&D-driven productivity improvements (and vice versa). Conversely, firms that experience an adverse shock in their access to imported input varieties have been found to suffer a severe hit to their productivity; see specifically Gopinath and Neiman (2014), who exploit micro data from Argentina's currency crisis of 2001-2002.

Given the strong evidence on positive selection into exporting, it stands to reason that firms that import inputs from abroad and become more productive as a result would find themselves in a better position to commence or expand their exporting activities as well. Several papers have indeed sought to uncover explicit evidence of such a link from access to imports on the one hand to improved export performance subsequently. Using Argentina's earlier trade liberalization episode in the mid-1990s, Bas (2012) finds that firms that saw greater input tariff reductions in turn had a higher probability to commence exporting. Firm-level studies using data from France (Bas and Strauss-Kahn, 2014) and from China (Feng et al., 2016) have also uncovered how greater access to imports is linked to expansions in firms' export value, their range of exported varieties, as well as to possible upgrading in the quality of the products they export.

From a methodological standpoint, the above body of micro empirical studies – that explore the effects of imported inputs on productivity, and possible subsequent links to export performance – has benefited from the ability to exploit country trade liberalization episodes in a firm-level data setting in order to design a credible identification strategy. To the extent that these events are either unanticipated or otherwise plausibly exogenous from the perspective of the firm, the variation across firms in the change in tariff exposure they experienced, often involving the use of a shift-share instrumental variable, has been leveraged to identify a causal effect of improvements in access to imports on firm-level outcomes.

We conclude this review of evidence on firm-level importing and exporting by highlighting some broader limitations of this literature's ability to shed more detailed light on GVC activity. Though such empirical work provides an understanding of firm-level trade patterns for a single country, this by definition only constitutes a particular snapshot of GVCs that pass through the country, but not information about the input-output links along an entire GVC. While there is work in this vein that seeks to dive deep to map out an entire GVC, this is typically limited to case studies or teardown analyses of particular goods such as the iPod (Dedrick et al., 2010) or bicycles (World Bank, 2020, Chapter 1). One can envision that there may be more systematic data efforts in the future that seek to unpack the complete GVC for multiple firms in a given industry, particularly one in which the set of inputs that lie upstream in the value chain is easy to enumerate and identify, in order to uncover a more comprehensive set of empirical facts on firm-level participation in GVCs.

---

[37] Colantone and Crinò (2014) document a relationship between imported inputs and an expansion in domestic products for a panel of European countries, albeit using industry-level rather than firm-level data.

3.2 Evidence on Buyer-Supplier Matching

Moving beyond data that reports on the import and export activities of individual firms, empirical work related to GVCs has benefited from the emergence of a limited number of datasets that contain information on firm-to-firm linkages, and even (in some data sources) firm-to-firm transaction flows. By identifying *both* parties that are connected in buyer-supplier or importer-exporter relationships, these provide the opportunity to examine the formation and features of particular links within domestic and global value chains. We discuss below the nature of such proprietary firm-to-firm datasets that have become available over the past decade, the key stylized facts on buyer-supplier matches that have been uncovered, as well as how these relate to the observations on selection into GVC participation discussed earlier in Section 3.1.

We touch first on studies based on datasets that contain firm-to-firm links in cross-border trade, this being the type of data that most directly speaks to flows within GVCs. An example of early work on this front is Blum et al. (2010), who performed a laborious merge of customs data from Chile and Argentina. Rather than merging datasets from different countries, several more recent studies have instead worked with customs data from individual countries that do a reasonable job of maintaining consistent records of the identity of the importing/exporting counterparty. Empirical researchers have successfully exploited such data from countries such as Norway (Bernard et al., 2018), Colombia (Eaton et al., 2016; Bernard et al., 2019), Costa Rica, Ecuador and Uruguay (Carballo et al., 2018), as well as France (Kramarz et al., 2020).[38]

These studies have unearthed a set of common empirical regularities. We overview these below, but would also direct interested readers to Section 2 of Bernard and Moxnes (2018) for a more comprehensive account of these facts on firm-to-firm links. Given an importer-exporter or buyer-supplier dataset with sufficiently universal sample coverage, one typically finds that the distributions of the number of importing firms per exporter (or buyers per seller) and the number of exporting firms per importer (or sellers per buyer) are very skewed. Most exporters sell to a small number of foreign buyers, and most importers purchase from a small number of foreign sellers. Only a select small subset of exporters successfully match and trade with a large number of foreign buyers – and similarly, only a small fraction of importers are linked with a large number of foreign sellers – even though these links tend to account for the bulk of bilateral trade value that is observed. Furthermore, at least one party in any given firm-to-firm match tends to be a large firm; there are few links observed between small buyers and small sellers.

Digging deeper, exporters that have few links (and are thus presumably smaller and less productive firms) tend on average to sell to importers that have many connections (and are thus presumably larger and more productive firms). On the other hand, exporters that have many links tend on average to be less selective, in that the set of importers that they sell to comprises firms with a lower average number of links. An analogous pattern applies in parallel to the number of exporters per importer (or sellers per buyer). Put in the vocabulary of graph theory, the set of

---

[38] See Section 3.3 for a discussion of how such data from the US and France have been exploited to establish facts about the duration of firm-to-firm relationships.

buyer-supplier links features “negative degree assortativity”.

Taken together, these empirical regularities are consistent with selection not just in importing or exporting *per se*, but in the formation of buyer-supplier matches. Such selection could be driven for instance by the presence of match-specific fixed costs (Bernard et al., 2018), so that it is larger and more productive firms that can incur and amortize such fixed costs, and that are therefore able to participate more extensively in domestic and global value chains. A number of recent studies have exploited shocks to trade patterns or trade policies to uncover adjustments in buyer-supplier matches that are broadly consistent with such selection mechanisms. Benguria (2021) shows that following the establishment of the U.S.-Colombia Free Trade Agreement, which lowered tariffs faced by U.S. firms in the Colombian market, U.S. exporters were more inclined to switch their firm-to-firm matches to establish links with larger Colombian importers. This is consistent with the higher profits from exporting to Colombia encouraging U.S. firms to incur the search costs to seek a better match in the local market. On the other hand, Sugita et al. (2020) uncovers how Mexican textile/apparel exporters became *de facto* less selective in their matching with U.S. importers following the removal of the Multi-Fiber Arrangement (MFA) quotas, which significantly expanded the pool of competing suppliers in the U.S. market from a major third-country source (China).[39]

Given the relatively limited number of datasets that record firm-to-firm flows in international trade, researchers interested in studying buyer-supplier matches have also turned to a range of other data sources that report only on links (and not actual transaction values) or sources that provide detailed information on domestic (but not necessarily cross-border) linkages. Atalay et al. (2011), for example, use Compustat to map out the production networks of U.S. publicly-listed firms; note though that this network mapping is a partial one, since the accounting rules only require these firms to disclose the identity of major customers that account for more than 10% of their revenue in a given reporting period. Researchers have also tapped into other commercial company-level datasets that identify firms’ supply chain partners, such as S&P Capital IQ (e.g., by Lim, 2018) and Factset Revere (e.g., by Huang et al., 2020; Amiti et al., 2020; Charoenwong et al., 2021). These other data sources are more comprehensive than Compustat, in that their coverage extends beyond publicly-listed firms, made possible by the respective data provider’s proprietary methods for aggregating supply-chain information from company news, announcements, or reports/filings; while this falls short of being a universal administrative dataset, it can nevertheless be sufficient for empirical applications in which it is the largest firms that are of interest. For example, the three papers cited above which tap the Factset Revere dataset have used it to study the recent U.S.-China tariff war, to explore if there have been adverse consequences on firms propagated through supply chain links. That said, even as the coverage of such commercial databases has been rapidly expanding, these remain more limited in the information they currently provide on actual input flows from suppliers to firms on the intensive margin.

The empirical literature on firm-to-firm networks has worked with two further classes of detailed

---

[39] Sugita et al. (2020) find at the same time that U.S. importers were conversely able to then match with higher-capability Mexican textile/apparel exporters, in line with the predictions of their model.

data. First, researchers have examined confidential data from credit bureaus or agencies that maintain records of the set of suppliers that individual firms are linked with. Bernard et al. (2019), for instance, use the Tokyo Shoko Research database from Japan, and demonstrate how the extension of a high-speed passenger rail-line increased the number of buyer-supplier links between firms in southern Japan and the rest of the country; this is highly suggestive evidence on the importance of face-to-face meetings for establishing supply chain links, as the new rail-line affected passenger travel times but not cargo shipping times.[40] Second, there has been a growing trend towards the use of administrative data on value added tax (VAT) records, that provide a close to full profile (modulo what are usually mild data censoring rules) of buyer-supplier transactions within the formal economy. Work that exploits such VAT data from countries including Belgium (Bernard et al., 2019; Dhyne et al., 2020) and Chile (Huneeus, 2018) has found that many of the empirical regularities found in importer-exporter data are replicated in domestic buyer-supplier networks as well. This includes the highly-skewed nature of the distributions of buyers per seller and sellers per buyer, as well as the "negative degree assortativity" of buyer-seller links, suggesting that selection into value chain participation is relevant too in the formation of domestic supply chain linkages. Demir et al. (2021) document a further feature in buyer-supplier matching that is present at least in the detailed VAT data from Turkey: Skill-intensive firms tend to purchase inputs from more skill-intensive suppliers, which the authors relate to the former's use of higher-quality inputs in their production processes. Alfaro-Urena et al. (2020) use firm-to-firm transaction data from Costa Rica of this nature to trace out how domestic firms benefit from spillovers through backward linkages when they become suppliers to multinational firms. Although the focus in these VAT datasets on domestic transactions represents a limitation in the extent to which they can shed light on GVC activity, the frontier of empirical research here is (arguably) not that far removed from working with merged administrative datasets that contain both domestic and international transactions; see for example Adao et al. (2020) for one such ambitious effort that has already been put together for Ecuador.

We touch briefly on a series of papers that has investigated the nature of shock propagation (both forward and backward) through buyer-supplier linkages, as these speak to how such shocks can be transmitted through and impact GVCs.[41] This line of work can be characterized by their relatively clean empirical design, often exploiting natural disasters as sources of near-experimental variation in the severity of shocks that firms are exposed to. Barrot and Sauvagnat (2016) make use of the above-mentioned Compustat data to show that natural disasters that affect supplier firms have adverse impacts on the market values of their major customers, particularly if the inputs provided are highly specific. Boehm et al. (2019) study how the 2011 Tohoku earthquake disrupted the supply chains of Japanese multinationals, as evidenced through the impact on the production of their affiliates located in the U.S. Carvalho et al. (2020) examine the aggregate implications

---

[40] See also Miyauchi (2019) who uses the same Japanese credit agency dataset to discipline a model of network formation that features agglomeration effects in buyer-supplier matching.

[41] This is related to the "bullwhip effect" in supply chain management: As shocks to demand for final goods work their way through supply chains, industries or stages of production that are further upstream experience larger variability in demand for their output. See Wang and Disney (2016) for a recent survey of the operations research literature on this topic.

of this same earthquake, focusing on the transmission of these shocks through Japan's domestic firm-to-firm network.[42] Using the FactSet Revere data on firms' suppliers, Kashiwagi et al. (2021) document how Hurricane Sandy adversely affected companies that were not directed affected by the disaster, but had links with other firms located in severely afflicted areas.

3.3 Evidence on the Relational Nature of GVC Links

The growing body of evidence on firm-to-firm links raises a natural question: What distinguishes buyer-supplier links within GVCs from regular market-based spot transactions? Toward this end, an active stream of work has spotlighted the *relational* nature of buyer-supplier links within GVCs. As emphasized by the World Development Report (World Bank, 2020), GVC links often exhibit durability or persistence over time, rather than merely being one-shot interactions. In this subsection, we review the progress that has been made in documenting the relational nature of buyer-supplier links in GVCs, in understanding the forces that support (or undermine) such links, as well as in exploring its consequences. Recent events such as the U.S.-China tariff war and the COVID-19 pandemic have raised interest in these issues, particularly as these can be informative of how durable GVC links might prove to be in the face of supply-chain disruptions.

It is proper first to acknowledge that economists are not alone in recognizing the importance of the relational nature of many GVC links. There is a large extant literature on value chain "governance" forms, that seeks to characterize and categorize the nature of buyer-supplier interactions within GVCs in more descriptive and textured terms, to which scholars in economic sociology, economic geography, and international business have contributed extensively. The seminal work of Gereffi (1994) and Gereffi (1999) laid out a key distinction between "producer-driven" and "buyer-driven" GVCs. The former ("producer-driven") refers to value chains in which the lead firm exercises a high degree of control over the production process, due for example to a desire to manage a proprietary technology; this often manifests itself in conscious decisions to vertically integrate production stages within firm ownership boundaries. The latter form ("buyer-driven") is instead characterized by input purchases that are conducted at arm's length, through market-based transactions. Gereffi et al. (2005) expanded on this typology, to introduce governance modes that lie between these two extremes, including in particular the notion of "relational value chains". Such value chain arrangements are more likely to emerge when the production process requires high levels of input customization or relationship-specific investment, and when the pool of available suppliers has sufficiently strong capabilities to deliver on these production requirements of the buyer. In such settings, the link between a buyer and its supplier is "built-up over time", with these repeated interactions often accompanied by an exchange of tacit knowledge. This generates a "mutual dependence" in the GVC link that makes "the costs of switching to new partners high" (Gereffi et al., 2005, p.86).[43]

---

[42] See also Todo et al. (2015) who find that Japanese firms in the earthquake area were better able to recover if they had supply chain links with domestic firms that were not directly affected.

[43] For completeness, the five value-chain governance modes in the Gereffi et al. (2005) typology are: hierarchical (or "producer-driven"), market (or "buyer-driven"), relational, modular, and captive.

The above characterization of "relational value chains" dovetails with a small but growing body of evidence uncovered from detailed micro data, on the average duration and durability of buyer-supplier links in international trade flows. The data that one needs in order to establish definitive facts on this front is fairly demanding, requiring time series data on cross-border transactions that reliably records the identities of both the buyer and seller. Such information on identities is in principle available in the raw customs paperwork that is filed in many countries, but in practice may require extensive work to clean of recording errors, as well as to pin down and link firm identities over time. Not surprisingly, the number of countries for which such evidence on the duration of firm-to-firm trade links is available remains small at present, although we anticipate that this will grow with more concerted data efforts.

Along these lines, Monarch and Schmidt-Eisenlohr (2017) document a series of stylized facts that speak to the duration of relationships between U.S. buyers and their foreign suppliers, using the U.S. Census Longitudinal Foreign Trade and Transaction Database (LFTTD).[44] Focusing on arm's-length imports at the HS-10 digit level by U.S. entities that are not in wholesale or retail, they find that new buyer-supplier links make up more than half of all new import relationships by count, but only for about a quarter of U.S. imports by value in 2011. Put otherwise, while there is dynamism in the formation of new links, it is ultimately existing or continuing relationships that account for a much greater share of trade value. Consistent with this, Monarch and Schmidt-Eisenlohr (2017) find that the value of transactions rises over time with the number of consecutive years that a buyer-supplier pair remain linked, while the hazard rate (that a link gets severed) falls steadily with relationship age. Eaton et al. (2014) uncover a similar pattern when zooming in on the subset of the LFTTD that comprises U.S. imports from Colombia, namely: a high initial hazard rate in the first year of a buyer-supplier link, but rising trade volumes conditional on survival.[45]

Using a particularly rich French customs dataset on firm-to-firm trade at the monthly frequency, Martin et al. (2020) illustrate that the duration of buyer-supplier links – defined as the number of months between when a French firm first imports a product from a given foreign supplier and when it first switches suppliers – has a distribution that features a long right tail. While close to 40% of links are one-shot transactions (i.e., with the buyer switching foreign suppliers in consecutive months), the median and mean durations are respectively 10 and 18 months, with a non-trivial share of around 5% of relationships sporting durations of seven or more years. The value transacted within a link is moreover increasing in the duration of the relationship. Building on these observations, Martin et al. (2020) develop a measure of "relationship-stickiness", which they construct by filtering the link duration data through the lens of a simple search model. This "relationship-stickiness" measure at the HS 6-digit level correlates positively with prior variables that have been associated with the

---

[44] The identity of the foreign supplier in U.S. import transactions is contained in the "manufacturer ID" variable on U.S. customs forms. This can be prone to inconsistencies in recording, but see Monarch and Kamal (2018) for a description of the data cleaning process they implemented to improve the accuracy of this variable.

[45] These expand upon the findings of Besedeš and Prusa (2006a, 2011), who document an analogous set of stylized facts using detailed product-level trade flows, though without exploiting confidential firm identifiers. Besedeš and Prusa (2006b) show that the survival probability of a given product-by-country import flow into the U.S. is higher for differentiated products, as classified by the Rauch (1999) measure.

formation of "relational value chains", such as product specificity (Rauch, 1999), contract-intensity (Nunn, 2007), and complexity (Hausmann and Hidalgo, 2014).

Focusing on U.S. imports from China in the LFTTD, Monarch (2020) reports that between 2002-2008, around one-half of U.S. buyer-supplier links in a given year persist into the next year. This survival rate is systematically higher for products that are more skill-intensive or contract-intensive (as measured by Nunn, 2007), in line with the idea that relationship-specificity in production tends to encourage the formation of relational GVCs. At the same time, U.S. buyers who switch a Chinese supplier are disproportionately more likely to select a new seller from the same Chinese city, which one can interpret as consistent with the costs of searching and gathering information being lower in a geographic location that is already familiar to the U.S. firm.

While the above empirical work has successfully exploited detailed micro data to deliver clear descriptive insights, it remains several steps removed from definitely decomposing the economic forces that account for cross-product variation in relationship duration. As has been touched on above, GVC links can exhibit persistence for a variety of reasons. Production of GVC inputs may require customization and relationship-specific investments, which may be more easily sustained through repeated interactions when the parties involved are sufficiently patient (Baker et al., 2002). GVC links could also be sticky because it is costly to search for alternative suppliers, or because there are barriers to the diffusion of information (Allen, 2014; Startz, 2016). On the other hand, the persistence of buyer-supplier links may itself depend on the size of the pool of potential buyers or suppliers, pointing to the relevance of considerations related to market power in the formation of GVCs. Understanding the relative importance of these various explanations for the relational nature of GVC links remains an interesting avenue for future work, which could moreover have practical consequences for the efficacy of policies intended to help a country's suppliers establish themselves in GVCs.

Moving beyond studies that have examined link duration, we round off this subsection by discussing empirical work that has studied relational contracting in global supply chains and their implications. In settings where contracts are inherently incomplete – either because not all contingencies can be contractually specified or because economic agents cannot count on the legal enforcement of all contractual terms – buyer-supplier links are instead often sustained by more informal or implicit relational arrangements. An insight that emerges from the relational contracting modeling framework of Baker et al. (2002) is that the organizational mode of "relational integration" – in which both the firm and supplier cooperate over multiple periods to deliver first-best effort levels, while conferring the firm greater residual rights of control – can be more readily maintained when the parties involved discount the future at a sufficiently low rate. While the intuition comes directly from the familiar Folk Theorem in repeated games, implementing a successful test of this prediction requires that one propose a credible proxy for the discount rates of GVC participants. Kukharskyy (2016) addresses this by drawing on a measure of cultural affinity towards "long-term orientation" from the social psychology literature (specifically, Hofstede et al., 2010); Kukharskyy (2016) shows that such measures of economic agents' "patience" indeed correlate positively with the propensity towards integration, as captured by both intrafirm trade shares and firm ownership shares.[46] This adds to a voluminous empirical literature exploring firm boundary decisions, and more specifically the propensity to integrate suppliers, that has been reviewed extensively in Antràs and Yeaple (2014) and Antràs (2016).

The empirical work on relational contracting has been enriched significantly by studies at the intersection of the fields of international trade and development, that have collected detailed information on transactions in specific supply chain links, often involving local suppliers and foreign buyers. By gathering fine-grained data and combining this with a keen appreciation of the institutional setting at hand, such studies have shed valuable light on the forces that can shape the terms of relational contracting arrangements, particularly in developing-country contexts where formal legal enforcement can be weak.

In such circumstances, repeated interactions can be important for establishing reputation when there is initial uncertainty about the seller's "type", and this appears to be key in determining the durability of a given buyer-supplier link. In the context of the Kenyan rose export market, Macchiavello and Morjaria (2015) find that the value of cut flowers sold by a domestic supplier to a foreign buyer tends to increase as the relationship ages, consistent with the seller developing a good reputation over time with the buyer. In response to a negative supply shock, Kenyan suppliers appear to prioritize which buyers to restart transactions with in a subtle and strategic manner that is consistent with their incentives to build reputation: Priority was placed on maintaining links with buyers with whom the supplier had a relationship of intermediate duration, rather than with new buyers (with whom the value of the relationship was smaller) or with long-standing buyers (who were already fully aware of the supplier's "type"). Using transaction-level data from Ecuadorian manufacturing supply chains, Brugués (2020) documents a pattern of rising quantities and falling prices in buyer-supplier transactions as the relationship between a given pair ages; given the prevalence of the use of trade credit as a terms of payment in this setting, Brugués (2020) rationalizes these patterns through a model in which suppliers use the implicit promise to maintain a link to incentivize buyers to make good on the trade credit extended to them. On the other hand, Startz (2016) shows how undertaking international travel to meet face-to-face with foreign suppliers plays an important role for Nigerian traders, when it is difficult for the foreign suppliers to credibly establish a public reputation. The relational nature of buyer-supplier links can also shape the payment and trade financing terms that exporters extend to importers, as shown by Antràs and Foley (2015).

While reputation can help to sustain buyer-supplier links, several studies have highlighted economic circumstances and forces that instead place relational supply chains under strain, particularly these improve the outside option of either the buyer or the seller. For example, Kenyan rose exporters are more prone to divert supply away from their regular buyers towards global auction markets instead when auction prices rise (Macchiavello and Morjaria, 2015). In the Rwanda coffee supply chain, buyers (coffee mills) are less likely to engage in relational contracts with sellers (coffee

---

[46] See Kamal and Tang (2014) for a similar effort at an empirical test using the LFTTD micro-data.

farmers) when they are situated near more competing coffee mills (Macchiavello and Morjaria, 2020).

When relational contracting becomes more difficult to sustain, buyers have been observed to opt instead for vertical integration of their suppliers, as has been the case in Costa Rica's coffee supply chain (Macchiavello and Miquel-Florensa, 2018). On the other hand, contrasting relational contracting with spot transactions in the Bangladesh garment supply market, Cajal-Grossi et al. (2020) find that relational buyers pay higher markups than spot buyers for products that are observationally equivalent on many dimensions. This is established through an original dataset with comprehensive information on material inputs, labor used and garment quality at the order level, and by exploiting within-supplier variation across different buyers. The higher markups paid by relational buyers can be viewed as an "efficiency wage" of sorts, to incentivize suppliers to meet benchmarks related to the reliable delivery of orders at short notice when necessary that can be hard to specify in an ex-ante contract.

4 Modeling GVCs: Macro Approaches

Having discussed empirical work in the last two sections, we next turn to overviewing theoretical work on the modeling of GVCs. We will begin this section by covering 'macro' approaches. As explained in the Introduction, we use this label to refer to work where the unit of analysis is the country-industry and where the emphasis is largely quantitative in nature. All the frameworks discussed below share the common feature of emphasizing the role of trade in intermediate inputs and of global intersectoral linkages in shaping the response of the world economy to various types of shocks, most notably trade policy shocks, but also productivity shocks, preference shocks, and labor supply shocks. Beyond providing a useful quantitative tool, these frameworks also provide a structural interpretation of the cells in a WIOT, and in some cases, these frameworks offer a microfoundation for the type of assumptions researchers implicitly make when invoking Input-Output analysis tools to compute the value-added content of trade flows or the positioning of countries in GVCs (as reviewed in Section 2). Although the vast majority of the papers reviewed in this section were written in the last ten years, we should of course acknowledge that quantitative work in international trade has a long tradition as a branch of the field of computable general equilibrium (or CGE) modeling (see Kehoe, 2005; Hertel, 2013; Hillberry and Hummels, 2013).

We will begin by reviewing the 'roundabout' model in Caliendo and Parro (2015), which has quickly become a benchmark model in the field. We will later discuss a few extensions of this framework, and we will close this section with an overview of recent work developing quantifiable models of multi-stage production.

4.1 Roundabout Models: The Caliendo-Parro Model

Caliendo and Parro (2015) present a multi-industry extension of the Eaton and Kortum (2002) Ricardian model of international trade. Although Eaton and Kortum (2002) already incorporated trade in intermediate inputs, it did so in a single-sector economy via a roundabout production assumption, and thus was not designed to assess the role of global interesectoral linkages in shaping the gains of international specialization and the response of the world economy to various types of shocks. Below, we review the basics of the Caliendo and Parro (2015) model, how it connects with the 'macro' measurement literature of GVCs reviewed above, and how this facilitates the computation of policy counterfactuals.

4.1.1 Theoretical Framework

Environment and Notation Caliendo and Parro (2015) consider a world with $J \geq 1$ countries and $S \geq 1$ sectors or industries. Producers in all sectors produce an output that can be interchangeably used as an intermediate input or as a final (consumer) good. There is a unique primitive factor of production, equipped labor, which is inelastically supplied in each country. All production technologies feature constant returns to scale and all producers behave competitively.

For the purposes of clarity, let us establish upfront the notation we use to index the economic variables in the exposition, which follows closely the one in Section 2, but deviates from that used in Eaton and Kortum (2002) and Caliendo and Parro (2015). We use *subscripts* i and j (and occasionally k) to refer to countries; whenever a pair of subscripts is used, the left subscript will denote the source country, while the right subscript will denote the destination country (so ij corresponds to a flow from i to j). On the other hand, we use the *superscripts* r and s to refer to industries; once again, whenever a pair of superscripts is used on a variable, the left superscript will be the identity of the source (i.e., selling) industry, while the right superscript will be the identity of the destination (i.e., buying) industry (so rs is a purchase from industry r by industry s).

Preferences and Technology The representative consumer in each country has preferences over the output of the S sectors given by

$$u(C_j) = \prod_{s=1}^{S} \left(C_j^s\right)^{\alpha_j^s}, \tag{13}$$

where C_j^s denotes consumption of a sector-s aggregate, C_j denotes the vector of the C_j^s's consumed in country j, α_j^s is the share of industry s in the expenditure of the country-j representative consumer, and $\sum_{s=1}^{S} \alpha_j^s = 1$.

Within each industry s, there is a continuum of varieties indexed by $\omega^s \in [0,1]$. Production of each variety is a Cobb-Douglas function of equipped labor, as well as intermediate inputs. More specifically, in country i, the production function for each industry-s variety is given by:

$$y_i^s(\omega^s) = z_i^s(\omega^s)\left(l_i^s(\omega^s)\right)^{1-\sum_{r=1}^{S}\gamma_i^{rs}} \prod_{r=1}^{S} \left(\mathcal{M}_i^{rs}(\omega^s)\right)^{\gamma_i^{rs}}. \tag{14}$$

Note that $\mathcal{M}_i^{rs}(\omega^s)$ is the amount of composite intermediates from industry r used in the production of variety ω^s in country i. The exponent γ_i^{rs} is the (constant) share of production costs spent on intermediate inputs from sector r by each industry-s producer in country i. It is assumed that $0 < \gamma_i^{rs} < 1$, and moreover that $0 < \sum_{r=1}^{S} \gamma_i^{rs} < 1$, so that the equipped labor share (or simply, value-added share) of production costs is strictly positive in all sectors and countries. The productivity shifter $z_i^s(\omega^s)$ is an i.i.d. draw from a Fréchet distribution with cumulative density function $F_i^s(z) = \exp\{-T_i^s z^{-\theta^s}\}$. The scale parameter T_i^s governs the state of technology of country i in industry s, while $\theta^s > 1$ governs (inversely) the dispersion of productivity in industry s across producers worldwide, thereby shaping comparative advantage.

The country-i composite in industry s, which is used both for final consumption (C_i^s), as well as to provide inputs to other sectors r ($\mathcal{M}_i^{sr}$), is a CES aggregate over the set of varieties on the unit interval:

$$Q_i^s = \left(\int q_i^s(\omega^s)^{(\sigma^s-1)/\sigma^s} d\omega^s \right)^{\sigma^s/(\sigma^s-1)}, \tag{15}$$

where $q_i^s(\omega^s)$ denotes the quantity of variety ω^s that is ultimately purchased, naturally from the lowest-cost source country. It is worth reiterating that the same CES aggregator over varieties applies to the industry-i composite, whether it is being consumed in final demand or being used as an intermediate input; we will return to this issue below.

Note that the framework captures the notion that countries not only import consumer goods, but also intermediate inputs from various industries and countries, with these imported inputs embodying *foreign value added*. Similarly, countries not only export consumer goods, but also intermediate inputs, thus generating *domestic value added* in production and exports of foreign countries. In sum, the Caliendo and Parro (2015) model captures important aspects of GVCs.

Equilibrium Consider the decision problem of either the representative consumer or a firm in country j, regarding which country to purchase variety ω^s from. As in Eaton and Kortum (2002), this corresponds to choosing the lowest-cost source country across $i \in \{1, \ldots, J\}$, after factoring in the unit production costs c_i^s and iceberg trade costs τ_{ij}^s across all potential source countries i.[47] The solution to this discrete choice problem yields an expression for the expenditure share of country j spent on industry-s varieties (intermediate or final goods) that come from country i:

$$\pi_{ij}^s = \frac{T_i^s (c_i^s \tau_{ij}^s)^{-\theta^s}}{\sum_{k=1}^{J} T_k^s (c_k^s \tau_{kj}^s)^{-\theta^s}}. \tag{16}$$

Country j's spending on country i-sector s's output is higher the higher the state of technology T_i^s, the lower the bundle cost c_i^s, and the lower the trade costs τ_{ij}^s associated with the i-s pair when selling in j. The unit production cost c_j^s is in turn obtained as the solution to the cost-minimization problem faced by each industry-s firm in country j, based on the production function (14). This is given by:

$$c_j^s = \Upsilon_j^s w_j^{1-\sum_{r=1}^{S} \gamma_j^{rs}} \prod_{r=1}^{S} \left(P_j^r \right)^{\gamma_j^{rs}}, \tag{17}$$

---

[47] We ignore tariffs and their implied tariff revenue, but they are modeled and taken into account in Caliendo and Parro (2015).

where Υ_j^s is a constant that depends only on the parameters γ_j^{rs}, and P_j^r is the ideal price index of the industry-r composite being used as an intermediate input in country j. Following Eaton and Kortum (2002), the expression for P_j^r is given explicitly by:

$$P_j^r = \kappa^r \left[\sum_{i=1}^{J} T_i^r \left(c_i^r \tau_{ij}^r \right)^{-\theta^r} \right]^{-1/\theta^r}, \tag{18}$$

where κ^r is a constant that depends only on σ^r and θ^r.[48]

Let X_{ij}^s denote the expenditure of country j on industry-s varieties from country i. This is the sum of country-j expenditures on the industry-s composite from country i, over both its use as an intermediate input and for final consumption. In turn, define: (i) $X_j^s = \sum_{i=1}^J X_{ij}^s$ as the total expenditure of country j on industry-s varieties; and (ii) Y_j^s as the value of gross output in industry s produced in country j. Having defined these objects, we can close the model by clearing the market for each industry in each country:

$$X_j^s = \sum_{r=1}^{S} \gamma_j^{sr} \underbrace{\sum_{i=1}^{J} X_i^r \pi_{ji}^r}_{Y_j^r} + \alpha_j^s \left(w_j L_j + D_j \right). \tag{19}$$

Note that the first term on the right-hand side of (19) is equal to the total purchases of intermediate inputs from industry s, where the sum is taken over all industries r that purchase intermediate inputs from s.[49] D_j is the national deficit of country j, computed as the sum of all sectoral and final-use imports of a country minus the sectoral and final-use outputs. Then, the second term on the right-hand side is the total purchases by country j on industry s for final consumption.

We finally impose trade balance, equating a country j's imports to its exports plus its observed deficit D_j:

$$\sum_{s=1}^{S} X_j^s = \sum_{s=1}^{S} \sum_{i=1}^{J} X_j^s \pi_{ij}^s = \sum_{s=1}^{S} \sum_{i=1}^{J} X_i^s \pi_{ji}^s + D_j \tag{20}$$

One can show that this last equilibrium condition can alternatively be derived from the equality of (equipped) labor income and total value-added.[50] The equilibrium of the model is then pinned down by the system of equations: (16), (17), (18), (19), and (20).[51]

---

[48] We assume that $\sigma^r < 1 + \theta^r$ for each r, in order for the ideal price index over this industry-r CES aggregate to be well-defined.

[49] The manipulation uses the fact that gross output of industry r in country j is equal to the world's total purchases from this country-industry.

[50] Aggregating (19) across sectors, and using (20), one obtains after some manipulations:

$$w_j L_j = \sum_{r=1}^{S} \left(1 - \sum_{s=1}^{S} \gamma_j^{sr} \right) \sum_{i=1}^{J} \pi_{ji}^r X_i^r = \sum_{r=1}^{S} \left(1 - \sum_{s=1}^{S} \gamma_j^{sr} \right) Y_j^r.$$

In words, the total wage payments to labor in country j are equal to total value-added across all sectors of j.

[51] Note that (16) comprises $J \times (J-1) \times S$ independent equations, since the shares π_{ij}^s need to sum to 1 for each j-s pair. Also, (17) and (18) each comprise $J \times S$ equations. The market-clearing condition (19) comprises $J \times S - 1$ independent equations, since one of these is redundant by Walras' Law. Finally, there are J trade balance conditions

4.1.2 Mapping the Model to Data

How does the Caliendo and Parro (2015) model map to available data from World Input-Output Tables (or WIOTs)? Remember from Section 2 (Figure 1) that a WIOT contains information on intermediate purchases by industry s in country j from sector r in country i, which we denote by Z_{ij}^{rs}. It also contains information on the final-use expenditure in each country j on goods/services originating from sector r in country i, which we denote by F_{ij}^{r}. Finally, the values of country-industry gross output Y_j^s and value-added V_j^s, as well as country-specific trade deficits D_j, can all be computed from the WIOT. It is then clear that the model offers transparent theoretical counterparts to all cells in a WIOT.

Furthermore, notice that the functional forms imposed by the model – namely Cobb-Douglas technologies, CES aggregators over varieties, and Fréchet distributions of productivity – imply that in line with the implicit assumption made in Section 2, the direct requirements $a_{ij}^{rs} = Z_{ij}^{rs}/Y_j^s$ are constant and given by:

$$a_{ij}^{rs} = \frac{\sum_{k=1}^{J} Z_{kj}^{rs}}{Y_j^s} \frac{Z_{ij}^{rs}}{\sum_{k=1}^{J} Z_{kj}^{rs}} = \frac{\gamma_j^{rs}}{\sum_{t=1}^{S} \gamma_j^{ts}} \frac{T_i^s (c_i^s \tau_{ij}^s)^{-\theta^s}}{\sum_{k=1}^{J} T_k^s (c_k^s \tau_{kj}^s)^{-\theta^s}}.$$

In other words, the model not only provides a structural interpretation of all the entries of a WIOT, but it also offers a microfoundation that rationalizes the type of matrix manipulations underlying the use of Input-Output Analysis to compute the value-added content of trade flows (for more on this, see Aichele and Heiland, 2018). This feature relates to the results in Burress (1994) demonstrating a similar homeomorphism in a closed-economy, input-output model with Cobb-Douglas technology and preferences.

It is worth stressing, however, a significant limitation of the Caliendo and Parro (2015) model in matching data from a WIOT. In particular, because production technologies and trade costs are common for inputs and final goods, and because preferences only vary across countries on account of sectoral spending shares, this framework imposes a unique market share of a given country i in the purchases of output of a given sector r by a destination country j, regardless of whether that output is designated for final-use or for use as an intermediate by other industries. In particular, note that the model imposes that for a given country pair ij and for a given sector r:

$$\frac{F_{ij}^{r}}{\sum_{k=1}^{J} F_{kj}^{r}} = \frac{Z_{ij}^{rs}}{\sum_{k=1}^{J} Z_{kj}^{rs}}$$

across all sectors s. To provide a concrete example, the model imposes that when buying finished products in the automobile industry (e.g., assembled cars), U.S. consumers spend their income across foreign sources of these finished goods in the same proportion in which U.S. auto makers buy

---

in (20). On the other hand, the equilibrium seeks to solve for the following objects: the shares π_{ij}^s (of which there are $J \times (J-1) \times S$ independent shares), the unit production costs c_j^s and price indices P_j^s (of which there are $J \times S$ each), as well as the $J-1$ wage levels w_j's (with one country's wage chosen as the numéraire) and the $J \times S$ expenditure levels X_j^s's. Thus, we have as many equilibrium conditions as variables to be solved for.

parts and components across foreign suppliers. As explained in Section 2.2, although traditional proportionality assumptions used to construct WIOTs tend to generate very similar if not identical trade shares for final goods and for intermediate inputs, more recent and sophisticated WIOTs have attempted to break these proportionality assumptions when constructing the tables. For instance, one of the contributions of the WIOD project was precisely to bring additional information to bear to distinguish imports across different end-use categories (see Dietzenbacher et al., 2013). Although even in state-of-the-art datasets, differences in trade shares are likely to be of little quantitative importance, we have argued in Section 2 that there are good reasons to believe (see de Gortari, 2019) that trade shares do vary significantly in the real world depending on *what* and *where* the input is used for, and we expect future WIOTs to more effectively exploit firm-level import and export data to document larger departures from the commonly-used proportionality assumptions.

A simple way to sidestep this issue, which Caliendo and Parro (2015) follow, is to simply aggregate intermediate inputs and final-good purchases as we have done when defining X_j^s in equation (19), and to map the trade share π_{ij}^s in the model to an empirical trade share computed as simply X_{ij}^s/X_j^s. As we shall see next, conditional on the structure of the model being the correct one (and thus ignoring the existing deviations of the model from actual WIOTs), it turns out that one can perform counterfactual analyses with knowledge of only these empirical trade shares, as well as a set of model parameters that can be easily calibrated to make the model fit the data exactly (namely, the Cobb-Douglas shares γ_j^{rs} and α_j^s) or that can be estimated (namely, the vector of trade elasticities θ^s).

4.1.3 Counterfactual Analysis: The Hat-Algebra Approach

As Caliendo and Parro (2015) note, the hat-algebra approach to counterfactual analyses devised by Dekle et al. (2008) in a one-sector model works equally well in their multi-sector model. More specifically, suppose one is interested in computing the counterfactual value of some key equilibrium variables of the model (such as real income per capita) following a shock to some of the parameters of the model. We denote the counterfactual value of a parameter or variable x with a prime (e.g., x') and use hats to denote the relative change in these variables, i.e., $\hat{x} = x'/x$. In practice, we will follow Caliendo and Parro (2015) in focusing on the effects of changes in trade costs τ_{ij}^s, though one could also use this approach to explore changes in the preference parameters α_j^s, or in the technology parameters T_i^r. For simplicity, assume that deficits D_j are held constant in the counterfactuals one studies.

Consider first the effects of trade cost shocks on trade shares. Using the hat algebra notation, it is easy to verify that (16) can be re-written

$$\hat{\pi}_{ij}^r = \left(\frac{\hat{c}_i^r \hat{\tau}_{ij}^r}{\hat{P}_j^r} \right)^{-\theta^r}. \tag{21}$$

In words, the percentage response of trade shares is purely shaped by the trade elasticity parameters θ^r and by the percentage shifts of the various trade cost parameters, as well as the percentage responses of the unit costs c_i^r, and the price index P_j^r. It is worth stressing that (21) is *not* an approximation: it holds exactly for any shock to trade costs, regardless of the size of the shock. Notice also, that the *level* of trade costs or the unobserved technological parameters T_i^r do not appear directly in these equations (though in some cases, it may be necessary to have knowledge of the initial level of trade costs to calibrate the relevant percentage change $\hat{\tau}_{ij}^r$ in these costs).

The responses of the unit costs c_i^r and the price index P_j^r to changes in the environment can be obtained from simple manipulations of equations (17) and (18). More specifically, plugging in the expressions for the trade shares from (16), we obtain:

$$\hat{c}_j^s = (\hat{w}_j)^{1-\sum_{r=1}^{S}\gamma_j^{rs}} \prod_{r=1}^{S} \left(\hat{P}_j^r\right)^{\gamma_j^{rs}}, \tag{22}$$

and

$$\hat{P}_j^r = \left[\sum_{i=1}^{J} \pi_{ij}^r \left(\hat{c}_i^r \hat{\tau}_{ij}^r\right)^{-\theta^r}\right]^{-1/\theta^r},. \tag{23}$$

There are two key features of these two sets of equations. First, the only variables in levels that appear in these equations are the trade shares prior to the shocks (which are observable), the Cobb-Douglas technological parameters γ_j^{rs} (which are retrievable from the data in a WIOT), and the trade elasticity parameters θ^r.[52] Second, it is clear from inspection that combining (22) and (23), one should be able to solve numerically for $\hat{c}_j^s$ and $\hat{P}_j^r$ as a function of these initial trade shares, as well as the percentage changes in wages ($\hat{w}_j$) and input trade costs ($\hat{\tau}_{ij}^r$). Plugging these resulting values of $\hat{c}_j^s$ and $\hat{P}_j^r$ into (21), this then allows us to express the changes in trade shares as a function of 'observables' (π_{ij}^s and γ_j^{rs}), the trade elasticity parameters θ^s, and the percentage changes in wages and trade costs.

We finally discuss how to trace the response of wages, as well as gross output and value-added, to the shocks. For that, we invoke the goods-market clearing conditions (19) and the trade balance conditions (20). In the counterfactual equilibrium, these can be re-written as

$$\left(X_j^s\right)' = \sum_{r=1}^{S} \gamma_j^{sr} \sum_{i=1}^{J} \left(\pi_{ji}^r\right)' (X_i^r)' + \alpha_j^s \left(\hat{w}_j w_j L_j + D_j\right) \tag{24}$$

and

$$\sum_{s=1}^{S} \left(X_j^s\right)' = \sum_{s=1}^{S}\sum_{i=1}^{J} \left(\pi_{ji}^s\right)' (X_i^s)' + D_j. \tag{25}$$

Noting that $\left(\pi_{ij}^r\right)' = \hat{\pi}_{ij}^r \cdot \pi_{ij}^r$, this system of equations delivers solutions for $\left(X_j^s\right)'$ and $\hat{w}_j$ as a function of changes in trade costs, observable pre-shock trade shares, and Cobb-Douglas parameters (as well as the elasticities).

---

[52] Specifically, if the model is not misspecified, γ_j^{rs} can be obtained by computing $\gamma_j^{rs} = \sum_{i=1}^{J} Z_{ij}^{rs}/Y_j^s$ for each country j and each pair of industries rs. Similarly, the Cobb-Douglas consumer spending shares α_j^s, which will appear in expression (24) below, can be obtained as $\alpha_j^s = \sum_{i=1}^{J} F_{ij}^s/\left(w_j L_j + D_j\right)$.

In sum, equations (21)-(25) demonstrate that in order to compute counterfactuals that shock trade costs while holding all other parameters constant, all that is required is the initial values of a set of variables that are easily retrieved from a WIOT, as well as values for the trade elasticities θ^s.

In their paper, Caliendo and Parro (2015) back out a log-linear estimating equation that amounts to running a triple-differences normalization of trade shares on a triple-differences normalization of asymmetric trade barriers such as tariffs.[53] This results in trade elasticities ranging from 0.37 for 'Other Transport' to 51.08 for 'Petroleum'. With these at hand and the other calibrated parameters, they quantify the consequences of NAFTA's tariff reductions. They find that although the model predicts that these tariff reductions increased intra-bloc trade very substantially (by 118% for Mexico, 11% for Canada, and 41% for the U.S.), the real income implications of NAFTA are predicted to be much more muted. In particular, according to the model, these tariff reductions increased real income in Mexico and the United States (by 1.31% and 0.08%, respectively) but decreased real income in Canada (by 0.06%).

Despite these meager effects, it is important to emphasize that the gains from trade-cost reductions in multi-sector models with intersectoral input linkages tend to be much larger than those implied in one-sector trade models, and are also larger than in multi-sector models without traded inputs (see Table 11 in Caliendo and Parro, 2015). This is further illustrated in Costinot and Rodríguez-Clare (2014), who compute the 'gains from trade' (or real income losses from moving to autarky) for various countries and for various possible environments. It turns out that these gains from trade are captured by the neat formula

$$GT_j = 1 - \prod_{s=1}^{S} \prod_{r=1}^{S} \left(\pi_{jj}^{s}\right)^{\widetilde{\gamma}_j^{rs} \alpha_j^s / \theta^s},$$

where remember that π_{jj}^s is the own trade share in sector j, α_j^s is the share of industry s in country j's final consumption, and θ^s is the sectoral trade elasticity. The new term $\widetilde{\gamma}_j^{rs}$ is the rs element of the Leontief inverse matrix $[\mathbf{I} - \mathbf{\Gamma}_j]^{-1}$, where $\mathbf{I}$ is the $S \times S$ identity matrix, and $\mathbf{\Gamma}_j$ is an $S \times S$ matrix with a typical element γ_j^{rs}. To illustrate the role of GVC linkages, it is particularly interesting to compare the estimates in columns 2 and 4 of their Table 4.1, which correspond to the gains from trade in a multi-sector competitive model with no input trade versus a multi-sector competitive model with input trade (as in Caliendo and Parro, 2015). These values are plotted in Figure 5. As is clear from the figure, gains from trade are higher for all countries in a world with intermediate input linkages, and the differences are quite large for some countries (the blue dotted line corresponds to a doubling of the gains from trade). On average, the gains from trade are 75% larger in a world of intermediate input trade.[54]

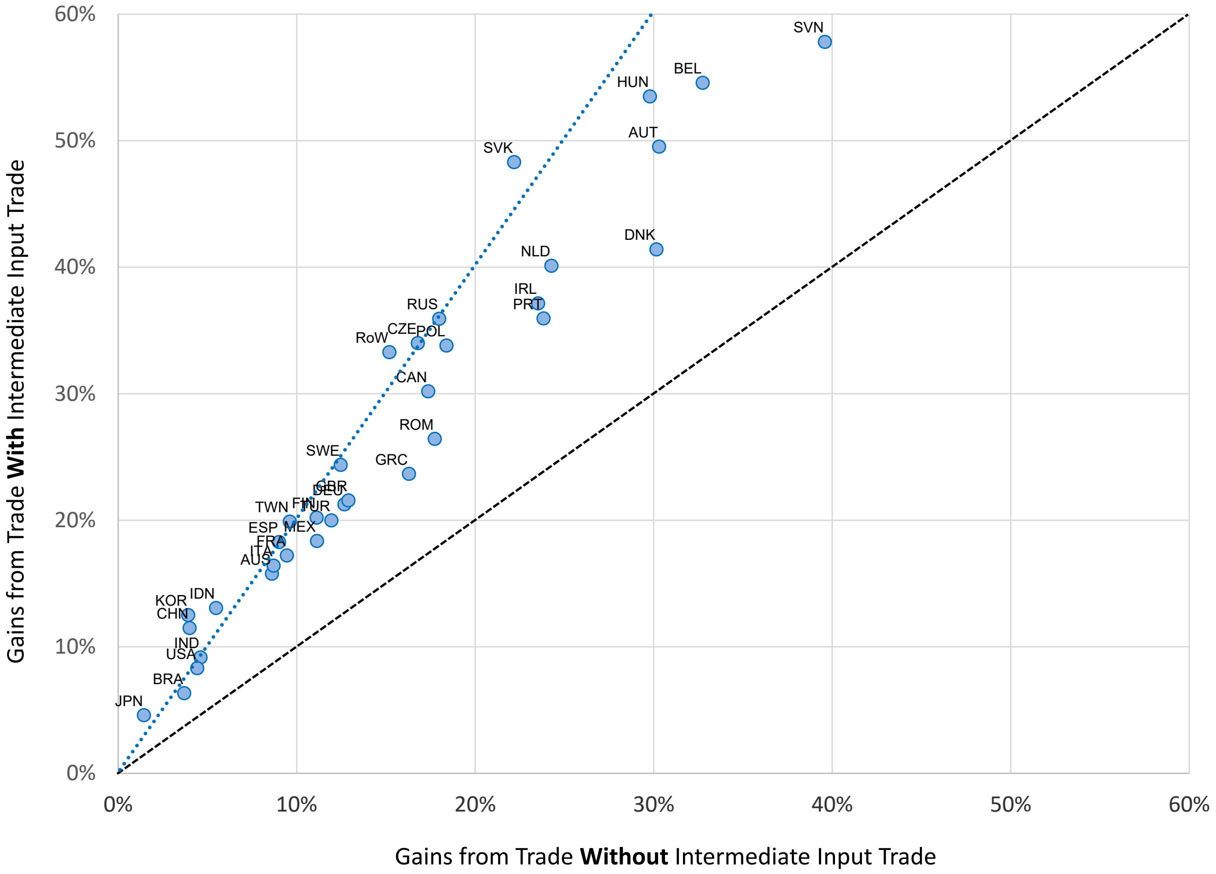
---

[53] More specifically, from equation (16), they obtain

$$\ln\left(\frac{\pi_{ij}^s}{\pi_{kj}^s}\frac{\pi_{ki}^s}{\pi_{ik}^s}\frac{\pi_{jk}^s}{\pi_{ji}^s}\right) = -\theta^s \ln\left(\frac{\tau_{ij}^s}{\tau_{kj}^s}\frac{\tau_{ki}^s}{\tau_{ik}^s}\frac{\tau_{jk}^s}{\tau_{ji}^s}\right).$$

[54] Giri et al. (2021) provide a more extensive analysis of the effects of several sources of sectoral heterogeneity for the size of the aggregate income gains from trade. An interesting observation in their work is that, because the

Figure 5: Gains from Trade With and Without Input Trade



4.1.4 Applications and Extensions

The last ten years have seen an explosion of quantitative work in the international trade field, and the awareness that intermediate input trade flows are a first-order feature of world trade, has led many researchers to adopt the Caliendo and Parro (2015) framework as the basis for conducting several types of counterfactual exercises. For instance, several authors have used the framework to quantify the effects of trade wars, and more specifically, of the recent U.S-China trade tensions (see, among others, Caceres et al., 2019; Beshkar and Lashkaripour, 2020; Ju et al., 2020; Charbonneau and Landry, 2018; Wicht, 2019). Another salient application is Dhingra et al. (2017)'s analysis of the aggregate income implications of UK's exit from the European Union (or Brexit). Other authors have employed the Caliendo-Parro framework to study the consequences of specific preferential trade agreements, such as the Transatlantic Trade and Investment Partnership (Aichele et al., 2016), or the U.S.-Japan Free Trade Agreement of 2019 (Walter, 2018). Furthermore, the framework has

---

relevant value of the trade elasticity in one-sector models should be lower than the sectoral trade elasticities estimated by Caliendo and Parro (2015), it is not hard to generate cases in which the gains from trade are actually *lower* with multiple sectors.

been employed to assess the economic consequences of the Belt and Road Initiative (De Soyres et al., 2018), and to quantify the welfare implications for Japan of productivity growth in emerging economies during the period 1995-2007 (Furusawa and Sugita, 2020). A more recent wave of work has employed the framework (or slight variants of it) to study the economic consequences of the ongoing COVID-19 pandemic, largely interpreting the shock as a labor supply shock (see Bonadio et al., 2020; Sforza and Steininger, 2020; Eppinger et al., 2020).

Beyond these largely "off-the-shelf" applications of the Caliendo and Parro (2015) model, it is worth discussing a bit more extensively a few more substantive extensions of their work. First, several sets of authors have worked to relax some of the assumptions of the framework so as to allow the model to more flexibly fit more of the cells in a WIOT. An early attempt is Alexander (2017), who develops a two-sector extension of Caliendo and Parro (2015) that delivers distinct trade shares for intermediate inputs and for final goods by allowing the technology parameter T_i^r to differ depending on whether producers in sector r produce final goods for consumers or inputs for other industries. More recently, Antràs and Chor (2019) have relaxed the assumption that iceberg trade costs τ_{ij}^s are only country-pair and (selling) industry specific. Instead, they consider the case in which trade costs are denoted by τ_{ij}^{rs} when goods/services in sector r from country i are shipped to industry s in country j (and they similarly denote by τ_{ij}^{rF} the trade costs incurred when goods/services in sector r from country i are shipped to final consumers in country j).[55] Antràs and Chor (2019) show that this simple extension allows the model to fully match *all* entries of a WIOT, and they also demonstrate that in order to conduct counterfactuals, one does *not* need more information than in the original Caliendo and Parro (2015), with 'hat-algebra' equations similar in nature to those developed in equations (21)-(25) above. This extension proves fruitful when studying how different shocks to the world economy shape the average positioning (upstream or downstream) of countries in GVCs.[56]

A second salient line of extensions have sought to relax the strong functional forms built into the Caliendo and Parro (2015) model. For instance, Caliendo et al. (2017) relax the Cobb-Douglas assumptions on preferences in (13) and on technology in (14), and show that this delivers an endogenous matrix of input-output or (direct requirement) links.[57] Baqaee and Farhi (2019) go even further and study nonparametric scenarios, while demonstrating that the gains from international integration can be significantly larger when input substitutabilities are lower than the unit ones imposed in Cobb-Douglas economies of the type studied by Caliendo and Parro (2015) (see also Fally and Sayre, 2018).

A third set of contributions has extended the model by introducing features from the economic

---

[55] This variation could reflect, for instance, underlying heterogeneity in the characteristics (weight, value, etc.) of the various inputs and final goods that are lumped together into a 'sector' in a WIOT. Naturally, it might also be driven by heterogeneity in the man-made trade barriers applied to these various industry subcategories. To the extent that different sectors buy different types of inputs in a given sector in different proportions, they will effectively face different trade costs, with the same being true of purchasers of final varieties.

[56] See Wicht (2020) for an exploration of the consequences of this same extended framework for the gains from trade.

[57] Ju et al. (2020) further relax the assumption of Cobb-Douglas technologies built into (14).

geography literature, such as the existence of multiple regions across countries, and the introduction of partial labor mobility across regions within countries. Two salient examples are Caliendo et al. (2018), who explore the implications of productivity shocks for the U.S. economy in the presence of interindustry and interregional intermediate input linkages, and Caliendo et al. (2019), who conduct a general-equilibrium analysis of the China trade shock, thus connecting with the reduced-form work of Autor et al. (2013) exploiting geographical variation across U.S. commuting zones in the incidence of import competition from China.

We close this section by briefly mentioning a few additional extensions of the Caliendo and Parro (2015) framework, which have permitted the quantitative evaluation of counterfactuals that were not feasible in the original model. Levchenko and Zhang (2016) present a dynamic model that allows them to trace the implications for trade flows and for real income of growth in sectoral total factor productivity in 72 countries and 19 sectors over 50 years. Di Giovanni et al. (2014) apply the same framework to isolate the implications of trade integration with China for other countries in the world economy. Caselli et al. (2020) incorporate sector-specific productivity shocks into the framework and study the extent to which international specialization increases or decreases the exposure of countries to productivity shocks in other countries. Morrow and Trefler (2017) explore another interesting extension of the framework which allows for multiple factors of production, and employ this framework to structurally study the implications of changes in trade costs, endowments and technology for the factor content of trade. More recently, Bagwell et al. (2018) have embedded a Caliendo and Parro (2015) in a model of international tariff negotiations, and study the counterfactual implications of tariff negotiations in the absence of the most-favored-nation clause. Finally, and even more recently, Rodríguez-Clare et al. (2020) have incorporated nominal rigidities into the framework to study the effects of trade shocks (and the China shock, in particular) on unemployment.

4.1.5 Critical Assessment

As the previous subsection has illustrated, the Caliendo and Parro (2015) model has quickly become a staple in the toolkit of international trade economists. It is important to close this section, however, with a critical assessment of the framework and its usage.

First, although the hat-algebra approach to counterfactual analysis is a remarkably useful tool, the minimal estimation requirements imposed by it are often overhyped. To be more precise, practitioners of this approach will often praise how parsimonious this approach is as opposed to the one followed in computable general equilibrium models, which instead requires estimating thousands of parameters. An often glossed fact, however, is that the hat-algebra approach requires the model to fit the data *exactly*, which amounts to calibrating all parameters of the models (or combinations of them) to values that ensure that the model fits the data exactly. In particular, this requires one to calibrate trade costs to infinity for the numerous country-sector to country-sector trade flows that take a value of zero in the data. Whether one should describe this approach as being ‘parsimonious’ is thus not entirely clear-cut.

Second, and more substantially, although quantitative work often requires strong assumptions on functional forms, calibrating thousands of parameters to fit the data exactly – even when those parameter values are implausible – can be problematic for the validity or reliability of the counterfactual predictions of those models. The problem is similar to overfitting in regression analysis leading to poor out-of-sample performance. As recently shown in Dingel and Tintelnot (2020), this is a particularly severe problem in spatial environments in which the data the model is fitted to contains a significant number of zeros, but note that even in the WIOD – a WIOT focusing on relatively rich countries – the share of zeroes is 13.7% in the matrix of input-use coefficients and 46.8% in the matrix of final-use column vectors.[58] It is then perhaps not too surprising that, as discussed in Kehoe et al. (2017), the performance of the Caliendo and Parro (2015) model in predicting the actual bilateral trade implications of NAFTA – as measured by the change in trade flows from 1991 to 2006 – is rather underwhelming. A natural counterargument is that 'many things' happened between 1991 and 2006, but the lack of 'external' evidence supporting the out-of-sample performance of these models remains problematic, and a clear area with room for improvement in future research.

4.2 Multi-Stage Approaches

The framework in Caliendo and Parro (2015) and its various extensions provide an interesting lens through which to interpret and quantify the implications of the rise in GVCs. The model features intermediate input trade and intersectoral linkages, and thus can be interpreted to capture the fact that global production takes places in "a series of stages with each stage adding value," to paraphrase our definition of GVCs in the Introduction. But by adopting a roundabout structure, the model essentially assumes that goods are produced via an endless sequence of steps, with each stage using inputs from prior stages in an infinite loop. Furthermore, all producers in a given sector use the same bundle of inputs in production, and operate the same technology in equation (14) regardless of the stage of production in which that production takes place.

In this section, we will turn attention to 'macro' approaches that specify multi-stage production technologies featuring a discrete number of vstages that add value in a pre-determined order. To simplify matters, we will begin by outlining a model with just two stages and no use of inputs or materials other than those proceeding from the prior stage, but will later incorporate the use of a composite bundle of inputs at each stage, as in roundabout models.

4.2.1 Two-Stage Case

Let us consider a simple multi-country Ricardian model of trade with multi-stage production inspired by the pioneering work of Yi (2003), and related to the frameworks in Yi (2010), Johnson and Moxnes (2019), and Antràs and de Gortari (2020). The world economy consists of $J \geq 1$ countries in

---

[58] Furthermore, there are a lot of observations that have very small value, and can thus only be rationalized with very large bilateral trade costs. As an example, the 25th percentile input-use coefficient value is 0.0001584 (compared to a mean value of 35.2).

which consumers derive utility from consuming differentiated varieties in a single sector. Preferences across varieties are CES, as in equation (15) in the Caliendo-Parro model reviewed in Section 4.1.

On the technology side, the good is produced combining two stages that need to be performed sequentially. Production in the initial stage $n = 1$ only uses labor, while the second stage of production combines labor with the good produced in the first stage. More specifically, we write these production technologies as follows:

$$y_i^1(\omega) = z_i^1(\omega^s) l_i^1(\omega^s) \tag{26}$$

and

$$y_i^2(\omega^s) = \left(z_i^2(\omega) l_i^2(\omega)\right)^{\alpha_2} \left(y_i^1(\omega)\right)^{1-\alpha_2}, \tag{27}$$

where $\alpha_2 \in (0, 1)$ denotes the labor share in stage-2 production, and $z_i^n(\omega)$ is labor productivity at stage n in country i. Firms are perfectly competitive and the optimal location $\ell(n) \in \{1, \ldots, J\}$ of the different stages $n \in \{1, 2\}$ of the value chain is dictated by cost minimization.

Countries differ in three key aspects: (i) their size, as reflected by the measure L_i of 'equipped' labor available for production in each country i (labor is inelastically supplied and commands a wage w_i), (ii) their geography, as captured by a $J \times J$ matrix of iceberg trade cost $\tau_{ij} \geq 1$, and (iii) their technological efficiency, as determined by the labor productivity terms $z_i^n(\omega)$. Following the lead of Eaton and Kortum (2002), we assume that $z_i^n(\omega)$ is drawn independently (across goods and stages) from a Fréchet distribution with cumulative distribution function $F_i^n(z) = \exp\{-T_i^n z^{-\theta^s}\}$.

Consider the lead firm problem of choosing the least-cost path of production to deliver consumption good ω to consumers in country j. Given equations (26) and (27), this amounts to choosing locations $\ell^j(1)$ and $\ell^j(2)$ to minimize

$$c\left(\ell^j(1), \ell^j(2)\right) = \tau_{\ell^j(2)j} \left(\frac{w_{\ell^j(2)}}{z_i^2(\omega)}\right)^{\alpha_2} \left(\frac{\tau_{\ell^j(1)\ell^j(2)} w_{\ell^j(1)}}{z_i^1(\omega)}\right)^{1-\alpha_2}. \tag{28}$$

Following the logic of Eaton and Kortum (2002), the hope is that the Fréchet assumption on the labor productivity $z_i^1(\omega)$ and $z_i^2(\omega)$ will deliver a convenient distribution for the equilibrium marginal cost of production of active GVCs, which will then facilitate a description of the general equilibrium of the model. Unfortunately, the minimum cost (28) associated with a given GVC path is *not* characterized by a particularly tractable distribution. The reason for this is that, although taking the minimum of a series of Fréchet draws is itself distributed Fréchet, and both $z_i^1(\omega)^{1-\alpha_2}$ and $z_i^2(\omega)^{\alpha_2}$ are Fréchet distributed, the product of Fréchet random variables is *not* Fréchet distributed.[59] As a result, papers adopting this lead firm approach to cost minimization with stage-specific Fréchet productivity draws need to resort to numerical methods to approximate the solution of their models, even when restricting the analysis to two-stage chains (see Yi, 2010; Johnson and Moxnes, 2019).

Antràs and de Gortari (2020) instead develop two alternative approaches that all permit a sharp

---

[59] Assuming a Leontief cost function (i.e., perfect complementarity) does not solve this problem because the sum of Fréchet random variables is not distributed Fréchet either.

and exact characterization of some of the features of the equilibrium, much as in the work of Eaton and Kortum (2002) and Caliendo and Parro (2015). The first approach considers a decentralized equilibrium in which stage-specific producers only minimize costs at their individual stage, and they do so with incomplete information about the productivity of certain suppliers upstream from them. More specifically, Antràs and de Gortari (2020) assume that firms know their productivity and that of the producers immediately upstream from them (i.e., their tier-one suppliers) when they commit to sourcing from a particular supplier, but they do not know the precise productivity of their suppliers' suppliers (i.e., tier-two suppliers, tier-three suppliers, and so on). Similarly, consumers (or retailers procuring goods on their behalf) know the productivity of the final-good producers that supply them goods, but not of the final-good producers' suppliers, and thus make purchase decision before knowing the actual minimum cost at which they will able to buy these goods.[60]

The second approach to gain tractability consists in simply treating the *overall* (i.e., chain-level) unit cost of production of a GVC flowing through a sequence of countries as a draw from a Fréchet random variable with a location parameter that is a function of the states of technology and wage levels of *all* countries involved in that GVC, as well as of the trade costs incurred in that chain.[61] To motivate this assumption, consider a given production path $\boldsymbol{\ell} = \{\ell(1), \ell(2)\} \in \mathcal{J}^2$, where $\mathcal{J}$ is the set of countries in the world. Its associated *chain-level* production cost is naturally a function of trade costs, composite factor costs and the state of technology of the various countries involved in the chain. Yet, two chains flowing across the same countries in the exact same order may not achieve the same overall productivity due to (unmodeled) idiosyncratic factors, such as compatibility problems, production delays, or simple mistakes.

More formally, and building on the cost function in (28), Antràs and de Gortari (2020) assume that the overall productivity of a given chain $\boldsymbol{\ell} = \{\ell(1), \ell(2)\}$ is characterized by

$$\Pr\left(\left(z_i^1(\omega)\right)^{1-\alpha_2}\left(z_i^2(\omega)\right)^{\alpha_2} \leq z\right) = \exp\left\{-z^{-\theta}\left(T^1_{\ell(1)}\right)^{1-\alpha_2}\left(T^2_{\ell(2)}\right)^{\alpha_2}\right\}, \tag{29}$$

which amounts to assuming that $\left(z_i^1(\omega)\right)^{1-\alpha_2}\left(z_i^2(\omega)\right)^{\alpha_2}$ is distributed Fréchet with a shape parameter given by θ, and a location parameter that is a function of the states of technology in all countries in the chain, as captured by $\left(T^1_{\ell(1)}\right)^{1-\alpha_2}\left(T^2_{\ell(2)}\right)^{\alpha_2}$. A direct implication of this assumption is that the unit cost associated with serving consumers in a given country j via a given chain $\boldsymbol{\ell}$ is also distributed Fréchet, which then allows one to readily invoke some key results from Eaton and Kortum (2002) to characterize equilibrium prices and the relative prevalence of different GVCs.

First, it is straightforward to verify that the share of country j's income spent on final goods

---

[60] With two stages ($N = 2$), this is the only relevant source of incomplete information, but as mentioned below, Antràs and de Gortari (2020) show that incomplete information in upstream stages allows them to easily generalize their framework to an $N > 2$ stage environment.

[61] Antràs and de Gortari (2020) also consider yet a third alternative decentralized approach inspired by the work of Oberfield (2018), in which technology is again specified at the stage level (rather than at the chain level), but in which productivity is buyer-seller specific. By appropriate choice of functional forms,they build on Oberfield (2018) to show that this formulation can also deliver a Fréchet distribution of productivity *at the chain level*.

produced under a particular GVC path $\boldsymbol{\ell} \in \mathcal{J}^2$ is given by

$$\pi_{\boldsymbol{\ell} j} = \frac{\left(\left(T^1_{\ell(1)}\right)^{\alpha_n}\left(\left(w_{\ell(1)}\right)^{\alpha_n}\tau_{\ell(1)\ell(2)}\right)^{-\theta}\right)^{1-\alpha_2} \times \left(T^2_{\ell(2)}\right)^{\alpha_2}\left(\left(w_{\ell(2)}\right)^{\alpha_2}\tau_{\ell(2)j}\right)^{-\theta}}{\sum_{\boldsymbol{\ell} \in \mathcal{J}^2}\left(\left(T^1_{\ell(1)}\right)^{\alpha_n}\left(\left(w_{\ell(1)}\right)^{\alpha_n}\tau_{\ell(1)\ell(2)}\right)^{-\theta}\right)^{1-\alpha_2} \times \left(T^2_{\ell(2)}\right)^{\alpha_2}\left(\left(w_{\ell(2)}\right)^{\alpha_2}\tau_{\ell(2)j}\right)^{-\theta}}, \tag{30}$$

and in addition, the exact ideal price index P_j in country j is given by

$$P_j = \kappa\left(\sum_{\boldsymbol{\ell} \in \mathcal{J}^2}\left(\left(T^1_{\ell(1)}\right)^{\alpha_n}\left(\left(w_{\ell(1)}\right)^{\alpha_n}\tau_{\ell(1)\ell(2)}\right)^{-\theta}\right)^{1-\alpha_2} \times \left(T^2_{\ell(2)}\right)^{\alpha_2}\left(\left(w_{\ell(2)}\right)^{\alpha_2}\tau_{\ell(2)j}\right)^{-\theta}\right)^{-1/\theta}, \tag{31}$$

where κ is a constant that depends only on σ and θ.[62]

A few observations regarding equations (30) and (31) are in order. First, these equations are closely related to those in the Eaton and Kortum (2002) and Caliendo and Parro (2015) frameworks. In fact, when $N = 1$, we necessarily have $\alpha_2 = 1$, and (30) and (31) reduce to equations (16) and (18) in our overview of the Caliendo and Parro (2015) model. Second, and quite intuitively, GVCs that involve countries with higher states of technology T_i or lower labor costs w_i will tend to feature disproportionately in production paths leading to consumption in j. Third, high trade costs penalize the participation of countries in GVCs, but their effect is more subtle than in models without multi-stage production. Notice, in particular, that the effect of trade costs 'compounds': if all trade costs in a particular GVC increase by 10%, this GVC's spending share decreases by $\theta\left(2-\alpha_2\right)$ percent, rather than θ in the roundabout model (see Yi, 2010, for more on this magnification effect). Another implication of this compounding effect is that, in choosing their optimal path of production, firms will be more concerned about reducing trade costs in relatively downstream stages than in relatively upstream stages, as reflected in the higher exponent for $\tau_{\ell(2)j}$ than for $\tau_{\ell(1)\ell(2)}$ in equation (30). As Antràs and de Gortari (2020) demonstrate, this feature of the model generates a centrality-downstreamness nexus by which, *ceteris paribus*, relatively more central countries will tend to gain comparative advantage and specialize in relatively downstream stages, a nexus for which they provide suggestive evidence.

Before completing the discussion of the equilibrium and implications of quantitative models with multi-stage production, let us briefly illustrate how the results above easily extend to a multi-stage environment.

4.2.2 Equilibrium with Multiple Stages

The bulk of 'macro' quantitative work on GVCs has focused on models of the type developed above with only two stages. There are various reasons for this focus (more on this below), but one of them is that in the absence of a tractable framework to pin down the relative prevalence of various GVCs, estimating models with more than two stages is highly complex. One of the advantages of

---

[62] For the price index to be well defined, one needs to impose $\sigma - 1 < \theta$.

the formulation of technology in Antràs and de Gortari (2020) is that their equilibrium equations naturally extend to an environment with an arbitrary number of stages N. More specifically, by specifying a Fréchet distribution of productivity at the chain level, or by making suitable assumptions about incomplete information regarding upstream suppliers, Antràs and de Gortari (2020) find that the share of country j's spending on final goods produced under a particular GVC path $\boldsymbol{\ell} = \{\ell(1), \ell(2), ..., \ell(N)\} \in \mathcal{J}^N$ is given by

$$\pi_{\boldsymbol{\ell}j} = \frac{\prod_{n=1}^{N-1} \left(\left(T^n_{\ell(n)}\right)^{\alpha_n} \left(\left(w_{\ell(n)}\right)^{\alpha_n} \tau_{\ell(n)\ell(n+1)} \right)^{-\theta} \right)^{\beta_n} \times \left(T^N_{\ell(N)}\right)^{\alpha_N} \left(\left(w_{\ell(N)}\right)^{\alpha_N} \tau_{\ell(N)j} \right)^{-\theta}}{\sum_{\boldsymbol{\ell}\in\mathcal{J}^N} \prod_{n=1}^{N-1} \left(\left(T^n_{\ell(n)}\right)^{\alpha_n} \left(\left(w_{\ell(n)}\right)^{\alpha_n} \tau_{\ell(n)\ell(n+1)} \right)^{-\theta} \right)^{\beta_n} \times \left(T^N_{\ell(N)}\right)^{\alpha_N} \left(\left(w_{\ell(N)}\right)^{\alpha_N} \tau_{\ell(N)j} \right)^{-\theta}}, \quad (32)$$

where α_n continues to denote the labor share in stage n, and where β_n is defined as $\beta_n \equiv \prod_{m=n+1}^{N} (1-\alpha_m)$. Notice that GVC shares continue to feature a magnified effect of trade costs as well as an increasing trade-cost elasticity as one moves to more and more downstream stages (since β_n is increasing in n). The price index P_j in country j is again a simple power function of the denominator in (32), exactly as in equation (31).

To solve for equilibrium wages, notice that for all GVCs, stage-n value added (or labor income) accounts for a share $\alpha_n\beta_n$ of the value of the finished good emanating from that GVC. Furthermore, total spending in any country j is given by $w_j L_j$, and the share of that spending by j going to GVCs in which country i is in position n is given by $\Pr(\Lambda^n_i, j) = \sum_{\boldsymbol{\ell}\in\Lambda^n_i} \pi_{\boldsymbol{\ell}j}$, where $\Lambda^n_i = \left\{\boldsymbol{\ell} \in \mathcal{J}^N \mid \ell(n) = i\right\}$ and $\pi_{\boldsymbol{\ell}j}$ is given in equation (32). It thus follows that the equilibrium wage vector is determined by the solution of the following system of equations

$$w_i L_i = \sum_{j\in\mathcal{J}} \sum_{n\in\mathcal{N}} \alpha_n\beta_n \times \Pr(\Lambda^n_i, j) \times w_j L_j. \quad (33)$$

The system of equations is nonlinear because $\Pr(\Lambda^n_i, j)$ is a nonlinear function of wages themselves, and of the vector $\boldsymbol{P}$, which is in turn a function of the vector of wages $\boldsymbol{w}$. When $N = 1$, we have that $\alpha_N\beta_N = 1$ and $\Pr(\Lambda^n_i, j) = \pi_{ij} = (\tau_{ij}c_i)^{-\theta} T^1_i / \sum_k (\tau_{kj}c_k k)^{-\theta} T^1_k$. The equilibrium then reduces to the general equilibrium in Eaton and Kortum (2002). Antràs and de Gortari (2020) derive a set of sufficient conditions that ensure that this solution exists and is unique for an arbitrary number of stages N.

Although the equilibrium is thus straightforward to compute, it is worth pointing out that with J country and N stages, there will be J^N active value chains for each destination country j. Hence, although the model can we analyzed for an arbitrary number of stages, in empirical applications, computational constraints are still likely to constrain how large N (or J) can be. We will return to related computational constraints in Section 5.

4.2.3 Gains from Trade

To study the real income implications of trade in this framework, it is useful to first consider a 'purely-domestic' value chain that performs all stages in a given country j to serve consumers in the same country j. Let us denote this domestic chain by $\boldsymbol{j} = (j, j, ..., j)$. Invoking equation (32), plugging the expression for the price index $P_j = \kappa \left(\Theta_j\right)^{-1/\theta}$, and rearranging, we find

$$\frac{w_j}{P_j} = \left(\kappa \left(\tau_{jj}\right)^{\sum_{n=1}^{N} \beta_n}\right)^{-1} \left(\frac{\prod_{n=1}^{N} \left(T_j^n\right)^{\alpha_n \beta_n}}{\pi_{\boldsymbol{jj}}}\right)^{1/\theta}. \tag{34}$$

This formula is analogous to the one that applies in the Eaton and Kortum (2002) framework and the wider class of models studied by Arkolakis et al. (2012). An important difference, however, is that $\pi_{\boldsymbol{jj}}$ is *not* the aggregate share of spending on domestic intermediate or final goods (which are readily available in input-output datasets), but rather the share of spending on goods that are produced *entirely* through domestic supply chains. As a result, the sufficient statistic approach advocated by Arkolakis et al. (2012) is not feasible in this setting, and one needs to estimate the model structurally to back out $\pi_{\boldsymbol{jj}}$ from available data. For a similar reason, the hat algebra approach to counterfactual analysis proposed by Dekle et al. (2008) and implemented by Caliendo and Parro (2015) is not feasible in a multi-stage setting.

We will provide more details on the mapping between multi-stage models and data in the next subsection, but let us briefly overview some qualitative implications of the model for the aggregate income implications of trade shocks. First, notice that the share of spending $\pi_{\boldsymbol{jj}}$ on purely domestic chains will, other things equal, be lower, the larger is the number of stages, and thus the gains from trade emanating from multi-stage models are expected to be larger on this account. This result is similar to the one derived by Melitz and Redding (2014b) in an Armington framework with sequential production, and also bears some resemblance to Ossa (2015)'s argument that the gains from trade can be significantly larger in multi-sector models, with stages here playing the role of sectors in Ossa's framework.[63] In their estimated model, Antràs and de Gortari (2020) find that the gains from trade (i.e., the income losses from reverting to autarky) obtained in their model are much larger when $N = 2$ than those obtained from a version of their model without multiple stages (i.e., $N = 1$). These values are plotted in Figure 6 for the same set of countries considered by Costinot and Rodríguez-Clare (2014) and already plotted in Figure 5. As is evident

---

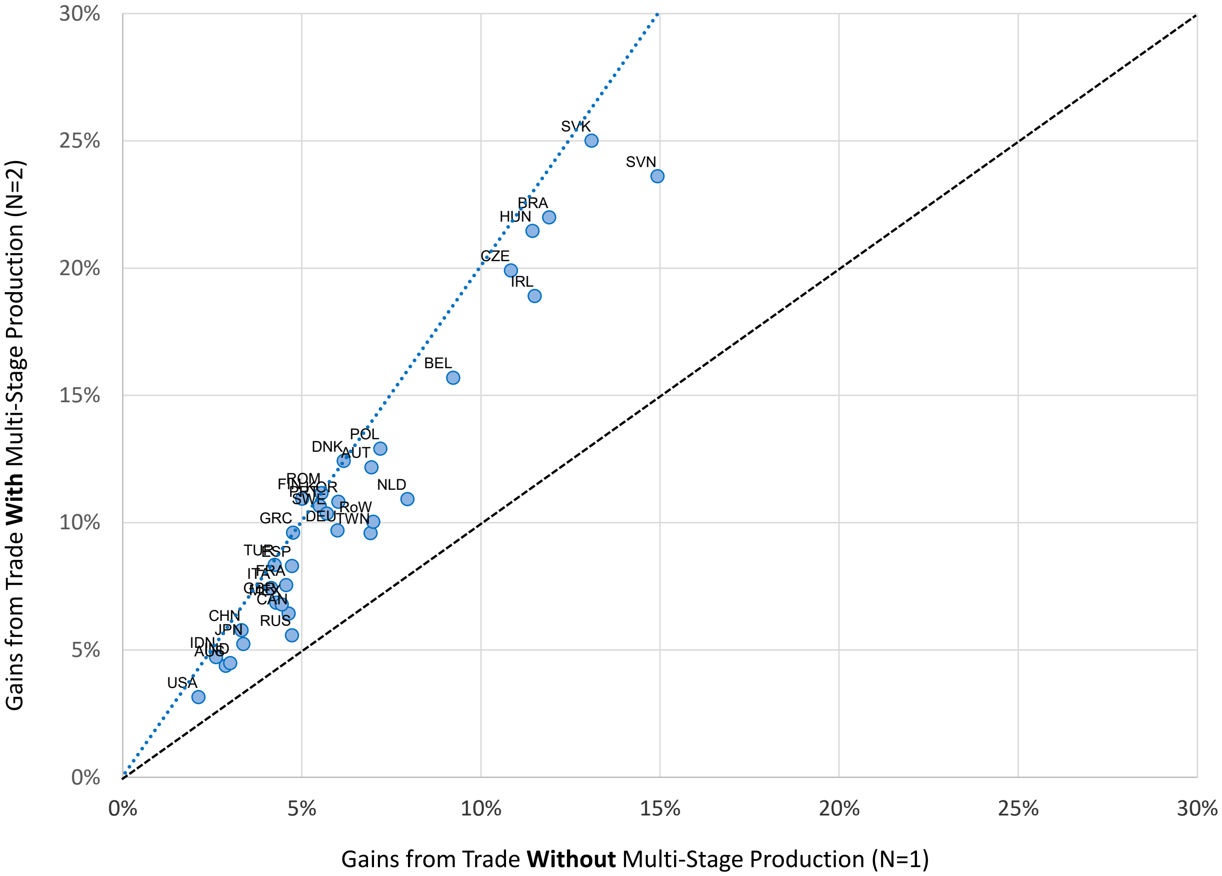
[63] For instance, Antràs and de Gortari (2020) show that if all countries are symmetric in all respects, and $\tau_{ij} = \tau$ for $i \neq j$ and $\tau_{ij} = 1$ for $i = j$, then the losses of reverting to autarky are given by

$$\hat{G}_j = 1 - \left(\prod_{n=1}^{N} \left(1 + (J-1)\,\tau^{-\theta\beta_n}\right)\right)^{-1/\theta}.$$

The gains from trade may thus become unboundedly large as production is sliced into more and more stages of production, i.e., $\lim_{N\to\infty} \hat{G}_j = 1$. Furthermore, it is straightforward to pick values for the value-added shares α_n such that the value-added to gross output ratio remains bounded even when $N \to \infty$.

from the figure, multi-stage production increase the gains from trade in all countries in a world with intermediate input linkages, and the differences are quite large for most countries (the blue dotted line corresponds to a doubling of the gains from trade). On average, the gains from trade are 72% larger in a world with multi-stage production.

Figure 6: Gains from Trade With and Without Multi-Stage Production



4.2.4 Extensions and Mapping to Data

Our discussion so far has been centered on a stylized model of multi-stage production, in which production processes are purely sequential. In this section, we briefly demonstrate the flexibility and applicability of this type of framework, and we show how they can easily nest both the Eaton and Kortum (2002) and Caliendo and Parro (2015) models. In the process, we also show how to map these multi-stage models to the observable data points contained in a WIOT.

A first straightforward extension is to allow production at each stage to use both 'equipped labor' as well as a bundle of intermediates or materials. Following Eaton and Kortum (2002), assume that this bundle is the same CES aggregator as in preferences. In other words, part of final-good production is not absorbed by consumers, but rather by firms that use those goods as a bundle of materials. Letting the cost c_i of the composite factor in country i be captured by a Cobb-Douglas aggregator, we have $c_i = (w_i)^{\gamma}(P_i)^{1-\gamma}$, where P_i is the ideal price index associated with preferences. As shown by Antràs and de Gortari (2020), all equilibrium equations – (30) through (33) – continue to hold with minor modifications, and the same is true about expression (34) for the gains from trade. Furthermore, when $N = 1$ the model reduces exactly to the Eaton and Kortum (2002) model.

How does one map this strict multi-stage generalization of the Eaton and Kortum (2002) model to the data? Although the 'GVC trade shares' in (30) are not observable in the data, it is straightforward to manipulate them to obtain closed-form expressions for various entries of a WIOT (when the data is collapsed into a single sector). Let us illustrate this for the case of the final-use vector. Notice that for final goods to flow from a given source country i to a given destination country j, it must be the case that country i is in position N in a chain serving consumers in country j. Defining the set of GVCs flowing through i at position n by $\Lambda_i^n \in \mathcal{J}^{N-1}$, the overall share of spending in country j in goods assembled in country i (i.e., in GVCs in which country i produces stage N) can be expressed as

$$\pi_{ij}^F = \frac{\sum\limits_{\ell\in\Lambda_i^N}\prod\limits_{n=1}^{N-1}\left(\left(T_{\ell(n)}^n\right)^{\alpha_n}\left(\left(c_{\ell(n)}\right)^{\alpha_n}\tau_{\ell(n)\ell(n+1)}\right)^{-\theta}\right)^{\beta_n}\times\left(T_i^N\right)^{\alpha_N}\left((c_i)^{\alpha_N}\tau_{ij}\right)^{-\theta}}{\Theta_j}, \tag{35}$$

where Θ_j is the denominator in equation (32). It then follows that final-good trade flows between any two countries i and j are then simply given by $\pi_{ij}^F \times w_j L_j$ (trade imbalances are ignored here but would be straightforward to incorporate). Computing intermediate input flows based on the 'GVC trade shares' in (30) is a bit more tedious, since one needs to take into account both vertical trade between two contiguous stages, but also intermediate input trade flows associated with the use of the bundle of inputs at each stage. Yet, as Antràs and de Gortari (2020) show, it is straightforward to obtain closed-form expressions for intermediate-input trade flows between any two countries i and j. With these expressions at hand, it then becomes feasible to estimate the key parameters of the model via maximum likelihood by minimizing the distance between various moments of a WIOT and their model counterparts.

The above framework can also be easily extended to a multi-industry environment that nests the Caliendo and Parro (2015) model. To see this, assume there are S industries indexed by $s \in \mathcal{S}$, with preferences given in (13), with sector specific Fréchet parameters θ^s, and with the cost of the bundle of labor and inputs used by country j in sector s given by

$$c_j^s = \Upsilon_j^s w_j^{1-\sum_{r=1}^S \gamma_j^{rs}} \prod_{r=1}^S \left(P_j^r\right)^{\gamma_i^{rs}},$$

as in equation (17) in the previous section. In such a case, letting $N = 1$, all equilibrium equations reduce exactly to the roundabout model of GVCs in Caliendo and Parro (2015). As shown in Antràs and de Gortari (2020) and de Gortari (2019), it is also straightforward to develop extensions of the framework that add multiple stages to certain variants of the Caliendo and Parro (2015) framework, such as those in Alexander (2017) and Antràs and Chor (2019), which allow certain parameters to be a function of the identity of producing country-industry pair, but also of the consuming country-industry pair.[64] A different matter is the ease with which these multi-industry extensions can be taken to the data, an issue we will address shortly.

We close this section by briefly mentioning some additional recent applications of quantifiable multi-stage models of GVCs. Beyond the work of Yi (2003, 2010), and Johnson and Moxnes (2019) mentioned above, another noteworthy recent contribution is Fally and Hillberry (2018). The main distinguishing feature of their framework is that they model production processes with a continuum of stages, and they consider the determination of the optimal set of stages that are carried out within firms and within countries. As in other papers in this literature, they use their model to quantify the importance of sequential production for the elasticity of trade flows to changes in trade frictions and for the aggregate income implications of changes in trade costs.

Other authors have recently applied or extended the framework in Antràs and de Gortari (2020) in interesting directions. Zhou (2020) uses a variant of the model to study the quantitative implications of the U.S.-China trade war, leveraging the fact that multi-stage frameworks allow for an independent analysis of the consequences of levying tariffs on inputs and on final goods. Lee and Yi (2018) develop a multi-factor extension of the same framework to study the interplay between the rise of GVCs and increased wage inequality. Finally, Yang (2018) incorporates within-country geography into the model to study the role of infrastructure in shaping the participation of countries in GVCs.

4.2.5 Critical Assessment

A feature of multi-stage models of GVCs that has hindered their widespread use is the fact that they are much harder to apply than roundabout models. In particular, even when the use of functional form assumptions dramatically simplifies the characterization of the equilibrium of the model, the fact that the hat-algebra approach cannot be used in this setting implies that the demands on estimation are much larger for this line of models. For instance, although a multi-sector model with multi-stage production can easily be written and solved, such a model is much harder to estimate because it requires estimating hundreds of parameters associated with technological parameters that can no longer be easily extracted from WIOTs. Having said this, and as we have argued in Section 4.1.5, the claim that roundabout models only require estimating a handful of parameters is somewhat disingenuous, as the hat-algebra approach essentially boils down to calibrating thousands of parameters to fit the data exactly. Naturally, this may lead one to rely too strongly on functional forms, which in turn can create problems of overfitting. We envision a future in which the estimation of macro models of GVCs will result in a more fruitful combination of calibration and estimation.

---

[64] de Gortari (2019) interprets these more flexible versions of the model as capturing specialized or customized inputs along GVCs.

5 Modeling GVCs: Micro Approaches

As we argued in Section 3, world trade flows are dominated by a small number of large firms that actively participate in GVCs and that capture large market shares in their sector's exports and imports. In more plain words, it is not countries or 'country-sectors' that participate in GVCs, but it is rather individual firms in those country-sectors that choose to do so. In this section we will review a body of work that has shed light on the decisions firms face over whether to participate in GVCs, and when designing their optimal GVC strategies.

We will proceed as follows. Initially, in Section 5.1, we will focus on decentralized frameworks in which firms make decisions only pertaining to the specific stages of GVCs in which they produce. A natural starting point is a model of selection into exporting in which an intermediate-input producer decides to participate in GVCs by exporting its output to foreign final-good producers instead of only providing inputs to domestic final-good producers. As will become apparent, under somewhat restrictive assumptions, this model of forward GVC participation will prove to be isomorphic to the seminal Melitz (2003) model of exporting. When these assumptions are relaxed, the fact that the exported goods are intermediate inputs rather than final goods carries significant implications. Next, we will consider a model of backward GVC participation, along the lines of Antràs et al. (2017), in which final-good producers select into importing, thus seeking parts and components from foreign suppliers rather than from local ones. Such a simple framework raises technical challenges and also generates a rich set of distinct implications relative to simple models of selection into exporting. Within the realm of decentralized models, we will finally consider frameworks in which both buyers and sellers make active participation decisions, thus connecting with the literature on firm-to-firm connections in trade.

In Section 5.2, we next turn to an overview of theoretical frameworks that consider the problem of a lead firm who chooses optimally the location of production of *all* the stages in a value chain, including stages that this (lead) entity does not directly participate in. This is motivated by the emphasis in the literature on the role of large multinationals in shaping the geography of GVCs. These lead-firm approaches raise novel technical hurdles, which the literature has circumvented in a variety of ways, as we outline in that section.

Finally, we close in Section 5.3 with an overview of work highlighting the 'relational nature' of GVCs, which is associated with a body of work emphasizing the distinctive role of search frictions, customized production and contractual insecurity in the recent rise of GVCs. This in turn opens the door for a succinct overview on the firm boundary and relational contracting decisions faced by agents participating in GVCs.

5.1 GVC Participation: Decentralized Approaches

5.1.1 Selection into Forward GVC Participation

Environment and Assumptions Consider a world consisting of J countries where consumers have preferences over a continuum of differentiated products. As in the models in section 4, preferences are CES and given by

$$Q_j^F = \left(\int_{\omega \in \Omega_j^F} q_j^F(\omega)^{(\sigma-1)/\sigma} \, d\omega \right)^{\sigma/(\sigma-1)} \tag{36}$$

in country j, where $\sigma > 1$ is the elasticity of substitution across varieties, and Ω_j^F is the set of consumption varieties available in country j. The resulting demand for variety ω in country j is $q_j^F(\omega) = E_j P_j^{\sigma-1} p_j^F(\omega)^{-\sigma}$, where E_j is total spending in country j, $p_j^F(\omega)$ is the price of variety ω, and P_j is the ideal price index associated with (36):

$$P_j = \left(\int_{\omega \in \Omega_j} p_j^F(\omega)^{1-\sigma} \, d\omega \right)^{1/(1-\sigma)}. \tag{37}$$

Consider next the supply side of the model. The only factor of production is equipped labor, which individuals in each country i supply inelastically in an amount L_i. Each final-good variety ω is produced by a single firm: the market structure in this final-good sector is characterized by monopolistic competition, and there is free entry into the industry. Production of final-good varieties employs 'equipped labor' to assemble a bundle of intermediates, much as in the 'macro' models reviewed in Section 4. The main novel feature here is that technologies will feature increasing returns to scale. More specifically, in order to produce, firms need to incur a fixed overhead cost equal to f_i^F units of labor. Unit costs in final-good production are in turn given by

$$c_i^F(\omega) = \frac{1}{z_i^F(\omega)} w_i^{1-\gamma} \left(P_i^I \right)^{\gamma}, \tag{38}$$

where P_i^I is the price index associated with the bundle of intermediates used in production, and is analogous to expression (17) in the Caliendo and Parro (2015) framework in Section 4. We assume that the bundle of inputs is a CES aggregator of a continuum of inputs, or

$$P_i^I = \left(\int_{\varpi \in \tilde{\Omega}_i} p_i^I(\varpi)^{(\rho-1)/\rho} d\varpi \right)^{\rho/(\rho-1)}, \tag{39}$$

where $\tilde{\Omega}_i$ is the set of input varieties available in country i, $p_i^I(\varpi)$ is the price paid in country i for input ϖ, and ρ governs the degree of substitutability across inputs. To simplify matters we will focus on the case in which all final-good producers share the same productivity $z_i^F(\omega) = z_i^F$ and in which trade costs on final goods across countries are purely ad valorem in nature and denoted by τ_{ij}^F when shipping from i to j.

Intermediate inputs are producing with only labor under a technology given by

$$f_i^I + q_i^I(\varpi) = z_i^I(\varpi) l_i^I(\varpi) \tag{40}$$

where f_i^I is an overhead cost incurred by suppliers and $z_i^I(\varpi)$ is labor productivity in input production. This upstream sector is also monopolistically competitive and there is free entry into the sector. Firms face an additional fixed cost of entry f_i^e incurred before their productivity level $z_i^I(\varpi)$ is drawn from some distribution $G\left(z^I\right)$. Intermediate inputs are tradeable across countries and incur iceberg trade costs τ_{ij}^I, but in addition, firms exporting inputs need to incur a fixed cost f_{ij}^X in order to export them from country i to market j. This last fixed cost will generate selection into exporting and thus into forward GVC participation.

Equilibrium with Nontradable Final Goods It is informative to consider first the case in which final goods are prohibitively to trade across countries ($\tau_{ij}^F \to \infty$). Consider the decisions of final-good producers in a given country j. Invoking constant-markup pricing, it is easy to verify that their profits are given by

$$\pi_j^F = \left(z_j^F\right)^{\sigma-1}\left(\left(w_j\right)^{\gamma}\left(P_j^I\right)^{1-\gamma}\right)^{-(\sigma-1)} B_j^F - w_j f_j^F, \tag{41}$$

where $B_j^F = \frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} w_j L_j P_j^{\sigma-1}$, and where we have imposed that, given free entry in both the downstream and upstream sectors, all income in all economies is labor income. Given the definition of the price index in (37) and the fact that firms are homogeneous, we obtain a simple expression for the measure of active final-good producers in country j:

$$N_j^F = \frac{L_j}{\sigma f_j^F}.$$

Note that each of these N_j^F producers will allocate a share $1-\gamma$ of their operating costs to purchasing intermediate inputs. Because unit costs are a constant multiple of operating profits, and the latter are brought down to $w_j f_j^F$ by free entry, we can conclude that intermediate input demand in country j is given by

$$P_j^I \mathcal{M}_j = N_j^F \times (\sigma-1)(1-\gamma) \times w_j f_j^F = \frac{\sigma-1}{\sigma}(1-\gamma) w_j L_j, \tag{42}$$

and is thus a simple multiple of aggregate income in market j.

We can now turn to the problem of an intermediate producer in country j. Notice that an intermediate input producer based in i selling to j will face a demand for the variety ϖ given by $q_j^I(\omega) = P_j^I \mathcal{M}_j \times \left(P_j^I\right)^{\rho-1}\left(p_j(\varpi)\right)^{-\rho}$. The profits obtained by this producer when exporting in country j are thus given by

$$\pi_{ij}^I\left(z_i^I\right) = \left(z_i^I\right)^{\sigma-1}\left(\tau_{ij}^I w_i\right)^{1-\sigma} B_j^I - w_i f_{ij}^X, \tag{43}$$

where $B_j^I = \frac{\sigma-1}{\sigma}(1-\gamma)\frac{1}{\rho}\left(\frac{\rho}{\rho-1}\right)^{1-\sigma} w_j L_j \left(P_j^I\right)^{\rho-1}$. It is then clear that behavior of intermediate input producers will be identical to that in the Melitz (2003) framework. In particular, only those firms from i with productivity $z_i^I \geq \tilde{z}_i^I$ will find it optimal to export to country j and thus forward-participate in GVCs in that destination country. This is in line with the empirical fact that exporters (regardless of whether they are final-good or intermediate-input producers) tend to be more productive than non-exporters (see Section 3.1). This selection into exporting will operate independently across countries, that is, the decision to export to a given destination j is not affected by this same firm's decisions in other markets. This feature greatly simplifies the construction of the general-equilibrium of the model, particularly when one assumes that the distribution $G\left(z^{I}\right)$ from which intermediate input producers draw their productivity is Pareto in all countries (see Chaney, 2008; Helpman et al., 2008).

In sum, when final-goods are nontradable, a model of forward GVC participation essentially reduces to a Melitz-framework of selection to exporting. Nevertheless, when final-goods are nontradable, the value added embodied in the exported intermediate inputs does not cross two borders, so it is not entirely clear that one should consider this a model of GVC participation. With that in mind, we next consider the case in which final goods are tradable.

Equilibrium with Tradable Final Goods Consider now the case in which trade costs associated with final goods are bounded. Notice first that conditional on a demand for intermediate inputs $P_j^I \mathcal{M}_j$ in country j, the behavior of individual intermediate input producers will be identical to that in the case with nontradable final goods. There will thus be again selection into GVC participation and entry decisions will be independent market-by-market. The main complication that arises once final goods are tradable is that the demand for intermediate inputs is harder to determine because it is not only a function of aggregate income in j, but also of aggregate income in other countries where final-good exporters sell. More specifically, profits for final-good producers in (41) now become

$$\pi_j^F = \left(z_j^F\right)^{\sigma-1} \left((w_j)^{\gamma} \left(P_j^I\right)^{1-\gamma}\right)^{-(\sigma-1)} \sum_{k \in \mathcal{J}} \tau_{jk}^F B_k^F - w_j f_j^F,$$

where $\mathcal{J}$ denotes the set of countries in the world, as in previous sections (given the absence of fixed costs of exporting, final-good producers export everywhere). Imposing free entry and noting that $B_j^F = \frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} w_j L_j P_j^{\sigma-1}$, produces a system of J equations that allows to solve for the measure of final-good producers N_j^F as a function of the vector of wages (w_j), market sizes (L_j) and parameters. Further imposing labor market clearing, allows one to solve for the vector of wages in terms of the parameters of the model. Noting that $P_j^I \mathcal{M}_j = N_j^F \times (\sigma-1)(1-\gamma) \times w_j f_j^F$, one can then compute intermediate input demand in country j. Because exported intermediate inputs get re-exported by final-good producers, this models does produce forward GVC participation in a strict sense.

The computation of this equilibrium is, however, involved and thus hard to characterize. Following Melitz (2003), it is useful to consider a world in which all J countries are symmetric and there is a unique level of final-good trade costs τ^F between country-pairs. In that case, $B_j^F = B_F$ for all $j \in J$, and it is easy to verify that intermediate input demand is given by the same expression as (42) above, although real wages are naturally higher in this variant of the model than in the one in which final goods are tradeable.

5.1.2 Selection into Backward GVC Participation

Environment and Assumptions We next turn to outlining a model of backward GVC participation which builds on the work of Antràs et al. (2017).[65] The framework again features a final-good (or downstream) sector and an intermediate-input (or upstream) sector. Consumer preferences over manufacturing goods are as in the forward GVC participation framework: individuals value the consumption of differentiated varieties according to Q_j^F in (36).[66]

Technology and market structure in the final-good sector is also largely analogous to that previous model. There exists a measure N_j^F of final-good producers in each country $j \in \mathcal{J}$, each producing a distinct differentiated variety ω, and the industry is characterized by monopolistic competition and there is free entry into the industry. Furthermore, production combines labor and intermediate inputs exactly as in the cost function in (38) and firms need to incur an overhead cost of production of f_i^F units of country i's labor before any production can occur. Unlike in the above model of forward GVC participation, we will now focus attention on a case in which labor efficiency z_i^F in downstream production is heterogeneous across producers and drawn from a continuous cumulative distribution $G_i\left(z_i^F\right)$ after incurring a fixed cost of entry equal to f_i^e units of labor. As in Melitz (2003), final-good producers only learn their productivity z_i^F after paying the entry cost, but are assumed to choose their sourcing strategy with knowledge of that core productivity level. The main novel assumption in this framework is that a firm from country i only acquires the capability to import intermediates inputs from a source country j after incurring a fixed cost equal to f_{ij}^M units of labor in country i (at a cost $w_i f_{ij}^M$). We denote by $\mathcal{J}_i\left(z_i^F\right) \subseteq \mathcal{J}$ the set of countries for which a firm based in i with productivity z_i^F has paid the associated fixed cost of offshoring, and often refer to $\mathcal{J}_i\left(z_i^F\right)$ as the *sourcing strategy* of a firm.

To emphasize the implications of selecting into importing, and thus of backward GVC participation, we follow Antràs et al. (2017) in assuming that the intermediate input sector is perfectly competitive with labor productivity differences across inputs and countries specified as in Eaton and Kortum (2002). More specifically, we adopt the technology in equation (40), but we set overhead costs $f_i^I = 0$ and assume that the value of z_i^I for a given location i is drawn (independently across locations and inputs) from a Fréchet distribution, $F_i(z) = \exp\{-T_i z^{-\theta}\}$. Beyond the fixed cost importers need to incur to purchase inputs from a given country, shipping intermediate inputs across countries also involves iceberg trade cost τ_{ij}^I.

Equilibrium with Nontradable Final Goods As in our discussion of the model of forward GVC participation, it again proves useful to first solve the model for the case in which final goods are prohibitively costly to trade across borders. As shown by Antràs et al. (2017), the Eaton-Kortum

---

[65] We will refrain from reviewing the vast literature on offshoring and global sourcing, which includes the work of Feenstra and Hanson (1996), Antràs et al. (2006), and Grossman and Rossi-Hansberg (2008), among many others. This work is surveyed in Chapter 2 of the 4th volume of this *Handbook* (Antràs and Yeaple, 2014).

[66] A significant difference is that tackling the general equilibrium of this type of models is computationally difficult, so Antràs et al. (2017) introduce an additional (non-manufacturing) sector that captures a large enough (constant) share of the economy's spending to ensure that the manufacturing sector faces a perfectly elastic supply of labor.

structure of the intermediate input market implies that conditional on a given global sourcing strategy $\mathcal{J}_i\left(z_i^F\right)$, the share of input purchases sourced from any country j by a firm from i with productivity z_i^F is given by

$$\chi_{ij}\left(z_i^F\right) = \frac{T_j\left(\tau_{ij}^I w_j\right)^{-\theta}}{\sum\limits_{k\in\mathcal{J}_i\left(z_i^F\right)} T_k\left(\tau_{ij}^I w_k\right)^{-\theta}}, \tag{44}$$

if $j \in \mathcal{J}_i\left(z_i^F\right)$, and by $\chi_{ij}\left(z_i^F\right) = 0$ if $j \notin \mathcal{J}_i\left(z_i^F\right)$. The numerator $T_j\left(\tau_{ij}w_j\right)^{-\theta}$ in equation (44) captures the *sourcing potential* of country j from the point of view of firms in i, while the denominator in this expression, which equals the sum of sourcing potentials of the countries included in a firm's sourcing strategy, summarizes the *sourcing capability* of that firm. We next note that the price index for intermediate inputs faced by a firm with productivity z_i^F can be expressed as

$$P_i^I\left(z_i^F\right) = \kappa\left(\sum_{k\in\mathcal{J}_i\left(z_i^F\right)} T_k\left(\tau_{ij}^I w_k\right)^{-\theta}\right)^{-1/\theta}, \tag{45}$$

where κ is a constant. Adding a new location to the set $\mathcal{J}_i\left(z_i^F\right)$ naturally increases the sourcing capability of the firm, and the increased competition across supplying sources leads to a lower price index paid by the firm for the bundle of inputs. Invoking constant-markup pricing and the price index for inputs in (45), one can then express the firm's profits conditional on a sourcing strategy $\mathcal{J}_i\left(z_i^F\right)$ as

$$\pi_i^F\left(z_i^F\right) = \left(z_i^F\right)^{\sigma-1}\left(w_i\right)^{-(\sigma-1)\gamma}\left(\kappa\sum_{k\in\mathcal{J}_i\left(z_i^F\right)} T_k\left(\tau_{ik}^I w_k\right)^{-\theta}\right)^{(\sigma-1)(1-\gamma)/\theta} B_i - w_i\sum_{k\in\mathcal{J}_i\left(z_i^F\right)} f_{ik}^M, \tag{46}$$

where B_i is a residual demand term that depends on aggregate spending on manufacturing goods, and the final-good price index in that sector.

As is clear from equation (46), when deciding whether to add a new country j to the set $\mathcal{J}_i\left(z_i^F\right)$, the firm trades off the reduction in costs associated with the inclusion of that country in the set $\mathcal{J}_i\left(z_i^F\right)$ – which increases the sourcing capability – against the payment of the additional fixed cost $w_i f_{ij}^M$. This tradeoff is similar to the one faced by exporters in our model of forward GVC participation (see equation (43)), who also trade off higher operating profits (via increased export revenue) versus higher fixed costs. Nevertheless, there is a very important difference between selecting into exporting and selecting into importing. In the former case, and given the standard assumption of constant marginal costs of production, the decision to service a given market is independent of that same decision in other markets. Conversely, in models of selection into importing, firms select into offshoring precisely to affect their marginal cost. As a result the marginal change in profits in equation (46) from adding a country to the firm's set $\mathcal{J}_i\left(z_i^F\right)$ depends on the set of other countries from which a firm imports, as well as those countries' characteristics. The problem of a firm optimally choosing its sourcing strategy is thus much harder to characterize, both analytically as well as quantitatively, since it requires solving a combinatorial problem with 2^J elements (where J is the number of countries). Despite these complications, Antràs et al. (2017) derive a series of theoretical results that facilitate a fruitful study of a multi-country model of backward GVC participation.

First, Antràs et al. (2017) note that given the fact that the profit function in (46) is log-supermodular in core productivity z_i^F and the firm's sourcing capability, no matter what the actual optimal set $\mathcal{J}_i\left(z_i^F\right)$ may be, more productive firms necessarily choose global sourcing strategies that give them (weakly) higher sourcing capabilities, which implies that their cost advantage is magnified by their sourcing decisions, thus generating an increased skewness in the size distribution of firms. Second, given the structure of the model, whether the decisions to source from different countries are complements or substitutes ends up depending only on the relative size of $(\sigma-1)(1-\gamma)$ and θ. Selection into importing features complementarity across markets whenever $(\sigma-1)(1-\gamma)>\theta$, that is, when (i) demand is relatively elastic (so profits are particularly responsive to variable cost reductions), (ii) inputs are relatively important in production (low γ), and (iii) input efficiency levels are relatively heterogeneous across markets (so that the reduction in expected costs achieved by adding an extra country in the set of active locations is relatively high). Third, whenever sourcing decisions are complementary, one can use standard tools from the monotone comparative statics literature to show that the sourcing strategies of firms follow a strict hierarchical structure in which the number of countries in a firm's sourcing strategy is (weakly) increasing in the firm's core productivity level, in line with empirical evidence. Fourth, in this same 'complements case', one can also show that, holding constant the market demand level B_i, a reduction in any trade friction (τ_{ij}^I or f_{ij}^M) leads to a (weak) increase in the set $\mathcal{J}_i\left(z_i^F\right)$ and also increases (weakly) firm-level bilateral input purchases from *all* countries. More specifically, the model predicts that a decrease in sourcing costs from China – perhaps due to a *China shock* applying to imported inputs – is expected to lead to an increase in firm-level U.S. intermediate input demand not only from China, but also from other sources including the U.S. itself, as long as one controls for demand conditions, a prediction for which Antràs et al. (2017) present reduced-form evidence.[67]

Beyond these comparative static results, Antràs et al. (2017) show that whenever global sourcing decisions are complements, one can adopt an iterative algorithm first proposed by Jia (2008), which uses lattice theory to greatly reduce the dimensionality of the firm's optimal sourcing strategy problem, a feature which in turns allows them to estimate and simulate the model with limited computing power.[68] In subsequent work, Arkolakis and Eckert (2017) have shown that a variant

---

[67] Antràs et al. (2017) also use the model to study how the aggregation of firms' sourcing decisions shapes aggregate input flows across countries, and show that their model nests several key workhorse trade models, such as the Eaton and Kortum (2002) and multi-country versions of the Melitz (2003) framework. Nevertheless, the model does not deliver a standard gravity equation for trade flows, but rather an *extended* gravity equation featuring third market effects.

[68] In plain words, the algorithm is based on the following two facts. First, that the addition of a country to a firm's global sourcing strategy $\mathcal{J}_i\left(z_i^I\right)$ will be profitable whenever the addition of that country to a subset $\mathcal{J}_i'\left(z_i^I\right)$ of $\mathcal{J}_i\left(z_i^I\right)$ is already profitable. And, second, that if the elimination of a country from a firm's global sourcing strategy $\mathcal{J}_i\left(z_i^I\right)$ is

of the same type of algorithm can be implemented to solve for the extensive margin of sourcing even when sourcing decisions are substitutes rather than complements, and Yang (2020) provides an application to an oligopolistic setting.

Although we have largely focused on the work by Antràs et al. (2017), the literature on importing and backward GVC participation is quite extensive. A few of the empirical papers demonstrating the productivity effects of global sourcing were reviewed in section 3.1. On the theoretical front, the work of Blaum et al. (2018) and Blaum et al. (2019) is also noteworthy. Blaum et al. (2018) build on the insights of Arkolakis et al. (2012) to provide sufficient statistics to measure the aggregate effects of input trade on consumer prices in an environment in which different firms may feature heterogeneous levels of involvement in global sourcing. Using the structure of their model, they show that firm-level data on value added and on domestic expenditure shares in materials is sufficient to compute the change in consumer prices due to a shock to the import environment. In Blaum et al. (2019), the same set of authors unveil that, inconsistently with the predictions of the Antràs et al. (2017) framework (see equation (44)), French firms tend to feature highly heterogeneous share of spending in inputs across a common set of origin countries (i.e., holding the sourcing strategy fixed). Furthermore, this variation is systematically related to firm size, with larger firms concentrating their import purchases in their top origin countries much more than smaller firms do. The authors then develop a model that rationalizes these patterns by incorporating vertical differentiation (i.e., variation in quality) in inputs, a complementarity between productivity and quality, and variation across countries in the ability to produce high-quality inputs.

Equilibrium with Tradable Final Goods We have seen so far that a model of selection into backward GVC participation produces much richer predictions than a baseline model of forward GVC participation. Strictly speaking, however, the fact that the framework above focuses on a model with nontradable final goods implies that goods never cross two borders, so again it is not obvious that it captures GVC participation. Fortunately, extending the model to the case of tradable final goods is straightforward, and also generates interesting insights.

Suppose then that trade in final-varieties is only partially costly and involves both iceberg trade costs τ_{ij}^{X} as well as fixed costs f_{ij}^{X} of exporting. Firm behavior conditional on a sourcing strategy is largely analogous to that above. In particular, after observing the realization of its supplier-specific productivity shocks, each final-good producer will continue to choose the location of production for each input to minimize costs, which will lead to the same marginal price index $P_i^I\left(z_i^F\right)$ for intermediate inputs obtained above in equation (45). The main novelty is that the firm will now produce output not only for the domestic market but also for a set of endogenously chosen foreign markets, which constitute the firm's 'exporting strategy' (denoted by $\mathcal{J}_i^X\left(z_i^F\right)$). We can then express the problem of determining the optimal exporting and sourcing strategies of a firm from

---

profitable, such elimination should also be profitable whenever the firm sources from a subset $\mathcal{J}_i'\left(z_i^I\right)$ of $\mathcal{J}_i\left(z_i^I\right)$.

country i with core productivity z_i^F as:

$$\begin{aligned}\pi_i^F\left(z_i^F\right) &= \left(z_i^F\right)^{\sigma-1}\left(w_i\right)^{-(\sigma-1)\gamma}\left(\kappa\sum_{k\in\mathcal{J}_i\left(z_i^F\right)}T_k\left(\tau_{ik}^I w_k\right)^{-\theta}\right)^{(\sigma-1)(1-\gamma)/\theta}\sum_{h\in\mathcal{J}_i^X\left(z_i^F\right)}\left(\tau_{ih}^X\right)^{1-\sigma}B_h \\ &\quad -w_i\sum_{k\in\mathcal{J}_i\left(z_i^F\right)}f_{ik}^M-w_i\sum_{h\in\mathcal{J}_i^X\left(z_i^F\right)}f_{ih}^X.\end{aligned}$$

Antràs et al. (2017) show that the nature of the interdependencies, as well as the theoretical results derived from them, continue to hold in this environment with active selection into both importing and exporting (see also Bernard et al., 2018). The key new feature of the above profit function is that it also exhibits increasing differences in any pair of export and import entry decisions. This has at least two implications. First, regardless of whether $(\sigma-1)(1-\gamma)\leq\theta$ or $(\sigma-1)(1-\gamma)>\theta$, any change in parameters that increases the sourcing capability of a firm – such as a reduction in any input trade cost τ_{ij}^I or f_{ij}^M, or an increase in any technology parameter T_j – will necessarily (weakly) increase the participation of the same firm in exporting. Second, restricting attention to the complements case $(\sigma-1)(1-\gamma)/\theta>1$, the model delivers a complementarity between the exporting and importing margins of firms. For instance, holding constant the vector of residual demand parameters B_i, reductions in the costs of trading final goods across countries will not only increase the participation of firms in export markets, but will also increase the number of countries from which a firm sources inputs. Furthermore, when $(\sigma-1)(1-\gamma)/\theta>1$, an increase in firm core productivity raises the firm's import *and* export participation by more than it would when one of these margins is shut down.

We close this section by outlining a series of extensions of this framework of backward GVC participation (some of these extensions will be discussed in more detail in Sections 5.1.3 and 5.3). To begin, it should be clear that it would be straightforward to follow the approach in Section 5.1.1 and recast the above framework such that the firm selecting into importing and exporting does not produce final goods, but rather intermediate inputs, which may themselves be re-exported to third countries. This would produce a framework in which a firm participates in GVCs both backwards *and* forward. It is also straightforward to reinterpret the sources of inputs in the Antràs et al. (2017) framework as regions rather than countries, so that the model can be applied to studying the formation of *domestic* production networks, as in the work of Bernard and Moxnes (2018) and Furusawa et al. (2017). As outlined in the next section, some authors have also used extensions of this framework to analyze how firms select into sourcing from particular suppliers rather than from particular locations (see Dhyne et al., 2020, for instance). Later in this survey, we will also review work – Antràs (2016) and Chor and Ma (2020), in particular – that develops incomplete-contracting extensions of the Antràs et al. (2017) framework, which permit an analysis of the extent to which backward GVC participation entails intrafirm or arm's-length intermediate input imports.[69] Hoang

---

[69] Relatedly, Carluccio and Bas (2015) use a variant of the framework to study the role of worker bargaining power in shaping the offshoring decisions of French firms.

(2020) has recently studied a dynamic version of the model in which the fixed costs of sourcing are sunk in nature, which leads to hysteresis in backward GVC participation, and she devises a partial identification approach to provide bounds on those sunk costs. Huang (2017) studies the implications of the model for how concentrated is importing in certain sources, and how that shapes the response of firm profitability to source-specific shocks (SARS in his empirical application), while Farrokhi (2020) applies the model to the crude oil industry to study the choice of suppliers refineries select and how much they buy from each. In recent work, Lu (2019) and Wang (2021) have also built on the framework to study the interdependencies between the sourcing decisions of firms and the profitability of innovation and automation, respectively. Finally, Laugesen (2018) and Fan et al. (2019) both study comparative statics in the version of the model with tradeable final goods, allowing the industry price index to adjust, while Fan et al. (2021) investigate the effects of input trade liberalization on the product mix of multi-product exporting firms.

5.1.3 Two-Sided Matching Frameworks

The micro approaches developed so far consider environments in which firms make operational decisions that lead them to select into export destinations or input sources. We next overview a body of work that instead considers frameworks in which firms match with other firms rather than matching with "countries". We will first summarize the work of Bernard et al. (2018), which constitutes a simple variant of the model of forward GVC participation in Section 5.1.1. We will next outline Bernard et al. (2018)'s extension of the Antràs et al. (2017) model of backward participation. We will finally quickly overview an alternative body of work that has explored the formation of GVCs using tools borrowed from network theory (cf. Jackson, 2010).

Models with Deterministic Matching Bernard et al. (2018) consider a framework that shares many features with the model of forward GVC participation developed in Section 5.1.1. As in that framework, there is a final-good sector, where a continuum of producers assembles consumer goods with heterogeneous productivity levels under increasing returns to scale and monopolistic competition, and an intermediate input sector that produces inputs with labor, also with increasing returns and under a monopolistically competitive market structure. Intermediate inputs are tradeable but shipping them across borders entails incurring both iceberg trade costs as well as fixed costs, which are paid by the exporter. The main innovation of Bernard et al. (2018)'s framework is that they interpret these exporting fixed costs as relationship-specific in nature, and thus they need to be incurred whenever attempting to reach out to a new costumer, even when a firm is already servicing other costumers in the same destination market.[70] More specifically, the profits an intermediate input producer from i with productivity z_i^I obtains when selling shipping inputs to a final-good producer in country j with productivity z_j^F is given by

---

[70] In Section 5.3, we will review a body of work that interprets some of these investments as reflecting search costs.

$$\pi_{ij}^{I}\left(z_i^I, z_j^F\right) = \left(\frac{\tau_{ij}^I w_i / z_i^I}{P_j^I\left(z_j^F\right)}\right)^{1-\sigma} P_j^I\left(z_j^F\right)\mathcal{M}_j\left(z_j^F\right) - w_j f_j^X,$$

where P_j^I denotes the price index for inputs – and $\mathcal{M}_j\left(z_j^F\right)$ the corresponding demand – associated with a final-good producer. These objects in turn depend on the extensive margin decisions of all intermediate input producers, which significantly complicate the analysis of the equilibrium. Nevertheless, because more productive final-good firms will tend to be larger and demand more inputs, a larger set of suppliers will optimally select into selling to them, and this will in turn reduce the price index faced by final-good producers and further boost their input demand. Bernard et al. (2018) solve for firm behavior and also for the industry equilibrium of the model whenever labor productivities upstream and downstream follow Pareto distributions. Among other results, Bernard et al. (2018) show that their framework predicts that both the distributions of buyers per exporter and of exporters per buyer are characterized by many firms with few connections and a few firms with many connections. Intuitively, large and productive suppliers will select into selling not just to large and productive final-good producers, but also to smaller and less productive buyers. Similarly, large and productive buyers will have many exporters (even less efficient ones) willing to sell to them. As already mentioned in section 3.2, Bernard et al. (2018) show that this assortative matching pattern is consistent with evidence from Norwegian transaction-level customs data from 2004–2012. They further show that the "buyer margin" of international trade explains a large fraction of the variation in aggregate trade. Finally, they also aggregate the model at the industry level, and show that it retains many of the properties of models of firm heterogeneity with CES preferences and a Pareto distribution of productivity, such as the Chaney (2008) model.[71]

We next outline a model of backward GVC participation which generates firm-to-firm transactions through the selection into importing decisions of final-good producers. The model is inspired by the theoretical framework in the working paper version of Dhyne et al. (2020), which in turn extends the framework in Antràs et al. (2017).[72] The main innovation is to interpret the fixed costs of sourcing as applying at the supplier level rather than at the country (or location) level. More specifically, suppose that final good producers can source inputs from particular suppliers in various countries only after incurring a fixed cost equal to $w_i f_{ik}^M$ units of labor, where i and k are the countries of the final-good producer and supplier respectively. Adding suppliers to the final-good firm's sourcing strategy is profitable because it lowers the price index of intermediate inputs in their cost function. Although, one could microfound via Eaton-Kortum-style assumptions why an increase in competition among suppliers reduces costs, it is simpler to just assume that the inputs produced by different suppliers are differentiated, regardless of the country in which they are produced. If the elasticity of substitution across suppliers' inputs is constant and given by $(1+\theta)/\theta$, this produces a

---

[71] Related work considering the "buyer margin" of trade includes Arkolakis (2010) and Carballo et al. (2018).

[72] See Bilgin (2020) for an alternative approach extending Antràs et al. (2017) to a firm-to-firm trade setting.

profit function for the final-good producer of the following type,

$$\pi_i^F\left(z_i^F\right) = \left(z_i^F\right)^{\sigma-1}(w_i)^{-(\sigma-1)\gamma}\left(\kappa\sum_{k\in J}\sum_{v\in\mathcal{Z}_k} z_k^I(v)\left(\tau_{ik}^I w_k\right)^{-\theta}\right)^{(\sigma-1)(1-\gamma)/\theta} B_i - w_i\sum_{k\in J}\sum_{v\in\mathcal{Z}_k} f_{ik}^M,$$

where $z_k^I(v)$ is a parameter governing labor productivity of supplier v in country k, and where the other parameters are as defined in our rendition of the Antràs et al. (2017) framework. Given this profit function, it is straightforward to see that the bulk of the results in Antràs et al. (2017) will continue to hold here. More productive final-good producers (with higher z_i^F) will optimally invest in (weakly) larger sourcing capabilities, and if $(\sigma-1)(1-\gamma)>\theta$, they will have sourcing strategies that involve a larger set of suppliers, with their marginal supplier being less efficient than the marginal supplier of a less productive final-good producer. By the same token, more efficient suppliers will be 'selected' by a larger share of final-good producers, and the marginal final-good producer will be less productive, the more productive is the supplier. These patterns very much resonate with the 'negative' assortative matching patterns produced by the Bernard et al. (2018) framework and unveiled in the empirical literature, as described in Section 3.2.

Although the two frameworks that we have outlined above feature firm-to-firm matching and trade, relationships are initiated by investments from only one party in the transaction. We are not aware of more 'symmetric' models of the type outlined above – with product differentiation, monopolistic competition and firm heterogeneity– in which both upstream and downstream (or both sellers and buyers) incur fixed costs to *deterministically* initiate relationships, although some models of search and matching, reviewed later in Section 5.3.3, often have that feature.[73]

Models with Stochastic Matching Moving beyond analyses of how bilateral pairs of trade relationships are deterministically created, there is a parallel literature that has adopted tools from the network theory literature to develop stochastic models of how firm-to-firm production networks are formed. Technology and market structure are quite distinct in many of these papers, but they all share the feature that pairs of producers are formed by 'chance', although the rate at which pairs form is sometimes driven by fundamental factors. An example of this type of model is the work of Eaton et al. (2018), who build on the tools from Eaton and Kortum (2002) to develop a model in which final-good producers get randomly matched with heterogeneous suppliers. Final-good producers have full bargaining power, so each input (or task) is bought at its unit cost, and only from the least-cost supplier of each input. When a final-good producer can produce an input (or perform a task) more cheaply than any other supplier, that task is not outsourced and is instead part of the firm's value added (value added and inputs are perfect substitutes in their setting, unlike all models we have reviewed so far). Eaton et al. (2018) ingeniously choose functional forms and productivity distributions to obtain a neat characterization of the general equilibrium of the model, which allows them to shed light on features of the labor share in French manufacturing, and how it

---

[73] To be clear, it is not obvious that allowing firms to select both into backward and forward GVC participation would invalidate any of the insights described above.

is shaped by trade integration.

Another related work is Oberfield (2018), who also provides a theory of the formation of firm-to-firm links in which random matching plays a key role. Oberfield (2018) consider a setting in which firms produce output combining labor with an input provided by another firm, according to a Cobb-Douglas production function. The productivity with which labor and the input are combined is buyer-seller specific (or *match*-specific) and characterized by a Pareto distribution with shape parameter θ. Furthermore, each final-good producer (or buyer) chooses the best match among a pool of potential suppliers, with the number of available potential suppliers characterized by a Poisson distribution. For the case of a closed economy in which all producers' output can, in principle, be used as an input by any other firm, Oberfield (2018) shows that this formulation delivers a Fréchet distribution labor productivity, just as in Eaton and Kortum (2002).[74] A key feature of Oberfield's framework is that particularly productive suppliers are likely to be employed by many firms, who in turn become more productive themselves by employing these highly productive suppliers. This feedback loop very much resonates with the mechanics of the Bernard et al. (2018) model described above. As Oberfield (2018) shows, an implication of these complementarities is that his model generates large differences in productivity and size across firms, and particularly so when the elasticity of output with respect to intermediate inputs is high.

Both the Eaton et al. (2018) and Oberfield (2018) frameworks are static in nature, but a growing number of papers have considered dynamic environments in which firm-to-firm links are shaped by randomness, but in which the stock of those links evolves over time. A pioneering study in a trade context is the work of Chaney (2014), who considers a model in which producers accumulate a set of customers or buyers over time. This process of link formation takes two forms. First, a firm meets new partners at random but in a way that is biased toward the location of the firm, with local matches more frequent than distant ones. Second, once a firm has acquired a network of distant contacts, it also acquires new customers as if it was producing from those locations. Chaney (2014) describes the dynamic evolution of firm-to-firm links, and shows that it is in line with several features of granular French firm-level export data. For instance, his framework predicts and the evidence confirms that the average squared distance of exports is a power function of firm size, and that the dynamic process of match formation generates path-dependence in a firm's export growth, along the lines of the empirical findings of Morales et al. (2019). In a follow-up paper, Chaney (2018) shows that a similar process of network formation can explain why trade flows tend to decline with distance with close to a unit elasticity.

Lim (2018) proposes an alternative framework for the dynamics of network formation. He first outlines a framework that is analogous to the model of exporting (or forward GVC participation) in Bernard et al. (2018) with two exceptions. First, he considers production technologies that feature the same degree of substitution between labor and inputs as across inputs, and second, as

---

[74] As mentioned in Section 4.2 (see footnote 61), Antràs and de Gortari (2020) show that in a multi-country environment with an arbitrary but finite number of stages, provided that the distribution of productivity in the initial stage ($n = 1$) is Fréchet distributed, this approach delivers a distribution for final-good prices that is also Fréchet distributed.

in Oberfield (2018), he assumes that firms' output can be indistinguishably sold to consumers or to other firms as inputs (so firms are neither exclusive final-good producers nor exclusive input producers). In order to be able to sell to other firms, firms need to incur fixed costs, and as in the frameworks above, larger and more productive firms will have more firm-to-firm links, both because they can better amortize the fixed costs of generating links, but also because other firms will find it profitable to pay the cost to set up links with them. Despite the complex nature of the model, Lim (2018) provides a neat characterization of its equilibrium in a closed economy. He then introduces dynamics of link formation by assuming that both firms' fundamentals (demand shifters and core productivity) as well as the costs of maintaining relationships follow a first-order Markov process. These features allow Lim (2018) to study the effects of business cycles on productivity, and the contribution of the entry and exit of firm-to-firm links for economic fluctuations. In more recent work, Huneeus (2018) presents an open-economy extension of the model in Lim (2018), which permits a study of the rich implications of international trade shocks – which are salient in his Chilean application – and how these shocks percolate along domestic production networks.

Overall, the body of work reviewed in this section is attempting to shift the focus from models in which firms make decisions about participating in GVCs in isolation, to environments in which firms' decisions and the shocks they face interact with each other, thus shaping the dynamics of economic activity and of aggregate trade flows in ways that are much richer than in environments without firm-to-firm links. The benefit of this approach is that it should result in more reliable quantitative and structural work, but this comes at the cost of a much greater complexity in analyzing and estimating these models.

5.2 Designing GVCs: The Lead-Firm Problem

Having outlined a number of decentralized approaches, we now turn attention to lead-firm approaches to the design of GVCs. The material in this section is motivated by the facts described in Section 3 indicating that world trade flows are crucially shaped by the operational decisions of a relatively small number of large firms. These 'superstar' firms do not only make exporting and importing decisions, but more generally, they design strategies to deliver their branded products to foreign consumers at the lowest possible cost. This leads them to seek suitable suppliers for the various stages in their value chains, and it also leads them to set up assembly plants in various countries to minimize the cost at which they make their goods available to distant consumers. In this section, we will outline a few variants of this lead-firm problem.

5.2.1 Multi-Stage Production

We begin by developing a simple model of firm behavior that formalizes the problem faced by a lead firm choosing the location of the various production stages involved in producing a consumer good. The good is produced combining N stages that need to be performed sequentially, and there are J countries in which consumers derive utility from consuming the good and in which the various stages can be produced. The last stage of production can be interpreted as final assembly and is indexed by N. As in previous sections, we will often denote the set of countries $\{1, ..., J\}$ by $\mathcal{J}$ and the set of production stages $\{1, ..., N\}$ by $\mathcal{N}$. At each stage $n > 1$, production combines a local composite factor with the good finished up to the previous stage $n-1$. Production in the initial stage $n = 1$ only uses the local composite factor.

Although the main insights of this section extend to more general specifications of technology, we will follow Antràs and de Gortari (2020) and focus throughout on Cobb-Douglas technologies at each stage. More precisely, we denote the unit cost of production of stage n in country $\ell(n)$ as

$$p^n_{\ell(n)}(\boldsymbol{\ell}) = \frac{1}{z_{\ell(n)}} \left(a^n_{\ell(n)} c^n_{\ell(n)}\right)^{\alpha_n} \left(p^{n-1}_{\ell(n-1)}(\boldsymbol{\ell})\, \tau_{\ell(n-1)\ell(n)}\right)^{1-\alpha_n}, \text{ for all } n \in \mathcal{N}, \tag{47}$$

where $\boldsymbol{\ell} = \{\ell(1), \ell(2), ..., \ell(N)\}$ is the path of production, $z_{\ell(n)}$ is a country-specific total factor productivity (TFP) term (common for all stages), $a^n_{\ell(n)}$ is the unit composite-factor requirement at stage n in country $\ell(n)$, $c^n_{\ell(n)}$ is the cost of the composite factor used at stage n in country $\ell(n)$, $\alpha_n \in (0,1)$ denotes the cost share of the composite factor at stage n, and $\tau_{\ell(n-1)\ell(n)}$ are iceberg trade costs associated with shipping goods from $\ell(n-1)$ to $\ell(n)$. Because the initial stage of production uses solely the local composite factor, we set $\alpha_1 = 1$.

Note that equation (47) also applies to the final assembly stage N, and a good completed in $\ell(N)$ after following the path $\boldsymbol{\ell}$ is available in any country j at a cost $p^F_j(\boldsymbol{\ell}) = p^N_{\ell(N)}(\boldsymbol{\ell})\, \tau_{\ell(N)j}$ (we use the superscript F to denote finished goods). For each country $j \in \mathcal{J}$, the goal of the firm is then to choose the optimal path of production $\boldsymbol{\ell}^j = \{\ell^j(1), \ell^j(2), ..., \ell^j(N)\} \in \mathcal{J}^N$ that minimizes the cost $p^F_j(\boldsymbol{\ell})$ of providing the good to consumers in that country j. The remainder of this section will seek to characterize the solution to this problem. The questions we will attempt to answer are: what forces shape the optimal assignment of stages to countries, and how do they exactly do so?

Free Trade: Comparative Advantage To begin, consider a simple variant of equation (47) in which we set all trade costs to 0, so $\tau_{ij} = 1$ for all $i, j \in \mathcal{J}$, and in which $z_{\ell(n)} = 1$, so productivity differences are purely shaped by the productivity of the composite factor at stage n. Iterating (47), the optimal path of production solves:

$$\boldsymbol{\ell}^j = \arg\min_{\boldsymbol{\ell} \in \mathcal{J}^N} p^F_j(\boldsymbol{\ell}) = \arg\min_{\boldsymbol{\ell} \in \mathcal{J}^N} \left\{ \prod_{n=1}^{N} \left(a^n_{\ell(n)} c^n_{\ell(n)}\right)^{\alpha_n \beta_n} \right\} \tag{48}$$

where

$$\beta_n \equiv \prod_{m=n+1}^{N} (1 - \alpha_m), \tag{49}$$

and where we use the convention $\prod_{m=N+1}^{N} (1 - \alpha_m) = 1$. In equation (48), while α_n is the cost share of the stage-n composite factor in stage-n production, $\alpha_n \beta_n$ is the cost share of this same stage-n composite factor in the *whole* global value chain (note that $\sum_{n=1}^{N} \alpha_n \beta_n = 1$).

As is clear from equation (48), we can break the cost-minimization problem into a sequence of N independent cost-minimization problems in which the optimal location of stage n is simply given by $\ell^j(n) = \arg\min_i \left\{ a^n_{\ell(n)} c^n_{\ell(n)} \right\}$, and is thus independent of the country of consumption j. It then becomes evident that the assignment of stages to countries is independent of the positioning of stages in the value chain and depends solely on standard relative production cost considerations, as in standard neoclassical trade theory. For instance, if the local composite factor is labor, there are a continuum N of stages, there are only two countries, and all firms in a given country share the same technology, the general equilibrium of the model becomes completely isomorphic to the celebrated Ricardian model of trade in Dornbusch et al. (1977), and thus the assignment of stages to countries is shaped by comparative advantage. Similarly, if $a^n_{\ell(n)} = 1$ for all n and $\ell(n)$, but the local composite factor combines capital and labor with different capital intensities in different stages, then the framework becomes related to the multi-stage neoclassical model in Dixit and Grossman (1982), and again whether countries specialize upstream or downstream depends on the interaction of their physical capital abundance and the factor intensity of various stages, regardless of the positioning of stages in the value chain.

Free Trade: Absolute Advantage To introduce a role for sequentiality, we now turn to a variant of (47) inspired by the work of Costinot et al. (2013), who in turn build on the insights from Kremer (1993). In particular, we now set $a^n_{\ell(n)} = 1$ and $c^n_{\ell(n)} = c_{\ell(n)}$ for all $n \in \mathcal{N}$ and $j \in \mathcal{J}$, but we allow TFP $z_{\ell(n)}$ to vary across countries. Standard trade theory would suggest that in the absence of comparative advantage differences across countries, the pattern of trade is indeterminate. Low TFP countries will face higher costs on account of their less efficient technologies, but in general equilibrium, their factor costs will adjust to equalize production costs across countries. Nevertheless, with sequential production matters are far less clear. In particular, iterating (47), the optimal path of production solves:

$$\boldsymbol{\ell}^j = \arg\min_{\boldsymbol{\ell} \in \mathcal{J}^N} p^F_j(\boldsymbol{\ell}) = \arg\min_{\boldsymbol{\ell} \in \mathcal{J}^N} \left\{ \prod_{n=1}^{N} \left(z_{\ell(n)} \right)^{-\beta_n} \left(c_{\ell(n)} \right)^{\alpha_n \beta_n} \right\} \tag{50}$$

where β_n is defined in equation (49). As in our case above, we can again break the cost-minimization stage by stage, and simply solve $\ell^j(n) = \arg\min_i \left\{ \left(z_{\ell(n)} \right)^{-1/\alpha_n} c_{\ell(n)} \right\}$. Note then that whether absolute TFP differences disproportionately affect upstream or downstream stages depends crucially on whether value-added intensity α_n rises or falls along GVCs. Costinot et al. (2013) develop a framework in which α_n effectively falls along the value chain, and thus conclude that absolute productivity differences across countries shape the specialization of countries in GVCs, with more efficient countries specializing in downstream stages of production. Although we will not solve for the general-equilibrium of the model, it should be clear that if we focus on the case in which $c_{\ell(n)} = w_{\ell(n)}$ (so the composite factor is just labor), there are a continuum of stages and only two countries, the model reduces to the Dornbusch et al. (1977) framework with a monotonic relative efficiency schedule that confers comparative advantage in downstream stages to high-TFP, rich countries.

Costly Trade We now consider an environment with costly trade, following the approach in Antràs and de Gortari (2020). For simplicity, we set $z_{\ell(n)} = a^n_{\ell(n)} = 1$ for all $n \in \mathcal{N}$ and $j \in \mathcal{J}$, and iterating (47), the lead-firm problem reduces to:

$$\boldsymbol{\ell}^j = \arg\min_{\boldsymbol{\ell} \in \mathcal{J}^N} p^F_j(\boldsymbol{\ell}) = \arg\min_{\boldsymbol{\ell} \in \mathcal{J}^N} \left\{ \prod_{n=1}^{N} \left(c^n_{\ell(n)}\right)^{\alpha_n \beta_n} \times \prod_{n=1}^{N-1} \left(\tau_{\ell(n)\ell(n+1)}\right)^{\beta_n} \times \tau_{\ell(N)j} \right\} \tag{51}$$

where β_n is again given in (49). Antràs and de Gortari (2020) emphasize two features of this problem. First, for general bilateral trade costs, a lead firm can no longer perform cost minimization independently stage by stage, and instead it needs to optimize over the whole path of production. Intuitively, the location $\ell(n)$ minimizing production costs $c^n_{\ell(n)}$ might not be part of a firm's optimal path if the optimal locations for stages $n-1$ and $n+1$ are sufficiently far from $\ell(n)$. A direct implication of this result is that the presence of arbitrary trade costs turns a problem of dimensionality $N \times J$ into J much more complex problems of dimensionality J^N each. As Antràs and de Gortari (2020) and Tyazhelnikov (2019) show, however, as long as technologies feature constant returns to scale, the *lead firm* can break the problem into a series of stage- and country-specific optimal sourcing problems, and then solve the problem via forward induction (starting in the most upstream stage), thereby solving the problem for all possible destinations j with just $J \times N \times J$ computations.

A second noteworthy aspect of the minimand in equation (51) is that the trade-cost elasticity of the unit cost of serving consumers in country j increases along the value chain. More specifically, note from equation (49) that, because $\alpha_n > 0$ for all n, we have $\beta_1 < \beta_2 < ... < \beta_N = 1$. The reason for this compounding effect of trade costs stems from the fact that the costs of transporting goods are (naturally) modeled as being proportional to the gross value of the good being transacted. Thus, as the value of the good rises along the value chain, so does the amount of resources used to transport the goods across locations. An implication of this compounding effect is that, in choosing their optimal path of production, firms will be more concerned about reducing trade costs in relatively downstream stages than in relatively upstream stages. Does this imply that more central countries that incur lower costs when importing and exporting should specialize in downstream stages? The answer to this question depends on the general equilibrium of the model, and if value-added intensity α_n rose sufficiently fast along the value chain, the fact that more central countries tend to command higher wages, might break the link between downstreamness and centrality, similarly to the countervailing general-equilibrium force in Costinot et al. (2013). Nevertheless, Antràs and de Gortari (2020) derive conditions under which a perfect correlation between centrality and downstreamness holds for *any* path of α_n, and they also provide exhaustive numerical analyses showing the robustness of the result in more general environments. Furthermore, building on the upstreamness measure in Antràs et al. (2012), and standard measures of centrality, they provide evidence for the existence of this relationship in the data.

We next briefly outline other work that has studied the optimal location of sequential production processes in the presence of trade costs. We have already mentioned above the work of Tyazhelnikov

(2019), which was developed independently from Antràs and de Gortari (2020). The work of Fally and Hillberry (2018) also incorporates trade costs, though in a more rudimentary manner, and is largely focused on delineating firm boundaries, so we will return to it in Section 5.3. Much more relevant is the earlier work of Harms et al. (2012) and Baldwin and Venables (2013) who both study two-country models in which the presence of costly trade in environments in which relative production costs (i.e., comparative advantage) does not rise or fall monotonically along production chains can generate interesting patterns of collocation, by which countries may end up capturing segments of value chains in which they are not particularly productive, but which take place soon before or soon after other stages in which they *are* disproportionately more productive.

Scale Economies Our discussion above has focused on the case in which technologies feature constant returns to scale. When solving for a lead-firm problem this is far from an innocuous assumption, especially in the presence of costly trade. It is fair to say that the literature has struggled to find a tractable way to incorporate increasing returns to scale and trade costs in models of sequential GVCs. We next illustrate some of the complications that have hindered progress and also report on some preliminary progress.

To build intuition, let us first consider a case with no trade costs and no TFP differences across countries. To model increasing returns to scale in the simplest possible manner, suppose that marginal costs continue to be independent of scale and given by the Cobb-Douglas technology in (47), but now assume that in order to activate country $j \in \mathcal{J}$ as a candidate location to produce stage n, the lead firm needs to incur a fixed cost equal to f_{ij} in terms of labor in the home country i of the lead firm. It should be clear that conditional on a subset of activated countries $\mathcal{J}_i \subseteq \mathcal{J}$, what one might call the lead firm's *GVC strategy*, the problem is analogous to that outlined in equation (48) except that the lead firm considers a smaller set of candidate location. Conditional on the profits obtained under alternative GVC strategies, the firm will then choose the strategy that delivers it the highest profit flow. In deciding whether to add a location, the firm will trade off the achieved marginal cost savings (and associated higher operating profits) with the upfront fixed cost. This trade off is quite similar to the one studied by Antràs et al. (2017) and overviewed in Section 5.1.2 above. In fact, if one assumes that the firm faces an isoleastic demand function, and the productivity terms $1/a^n_{\ell(n)}$ are drawn from a Fréchet distribution (independently across locations and inputs), $F_i(z) = \exp\{-T_i z^{-\theta}\}$, the resulting profit function for a given GVC strategy is given by

$$\pi_i^F\left(z_i^F\right) = \left(\kappa \sum_{k \in \mathcal{J}_i} T_k \left(c_k\right)^{-\theta}\right)^{(\sigma-1)/\theta} B_i - w_i \sum_{k \in \mathcal{J}_i} f_{ik},$$

where B_i is a residual demand term. It should be clear that this profit function is identical to that in equation (46) when $z_i^F = 1$, $\tau_{ij}^I = 1$ for all i, j, and $\gamma = 0$. The choice of the set of locations the lead firm activates can thus be studied using the exact same tools as developed in Antràs et al. (2017) (assuming, of course, that the set of activated countries is decided prior to the start of production).

The case with trade costs is, however, much harder to study. Intuitively in such a case, the *lead* *firm* problem cannot be solved independently for each destination market j, because whether a location ℓ constitutes a cost-minimizing location for stage n in a particular chain ending in j will be a function of the scale of this production node, and the latter is shaped by the overall level of production flowing through this node (potentially involving chains ending in destination markets other than j). As a result, dynamic programming ceases to be a powerful tool to simplify the problem (see de Gortari, 2020, for more details, and for an attempt to circumvent these complications).

Another approach to illustrate the complications that arise from the interaction of multi-stage production, trade costs and scale economies is to study a stylized general equilibrium model in which all firms produce a homogeneous good that requires N stages that can be produced in any of $J = N$ countries. Furthermore, assume the same sequential cost function as in (47), but now assume $z_{\ell(n)} = 1$ and $a^n_{\ell(n)} = (L^n_i)^{-\phi}$, where ϕ captures the role of external economies. Antràs and de Gortari (2016) study this environment and assume that ϕ is large enough to ensure a complete specialization equilibrium in which each stage is produced in exactly one country. Assuming logarithmic preferences and solving for the assignment of stages to countries that maximizes utilitarian world welfare, Antràs and de Gortari (2016) show that this problem solves

$$\boldsymbol{\ell} = \{\ell(1), \ell(2), ..., \ell(N)\} = \arg\min \sum_{i=1}^{N} \ln \tau_{\ell(N)i} + \sum_{n=1}^{N-1} \beta_n \ln \tau_{\ell(n)\ell(n+1)}, \tag{52}$$

where β_n is defined in (49). Intuitively, the optimal sequencing of production will simply seek to minimize the trade costs associated with the production process traveling through each of the J countries, 'visiting' each country exactly one time, and then returning to all countries in the form of a finished product. Antràs and de Gortari (2016) draw a connection between the optimization problem in (52) and the minimal distance Hamiltonian path problem in graph theory, or the associated travelling salesman problem (TSP) in combinatorial optimization. It is well known that both of these problems are NP-hard as they entail picking an optimal sequencing out of the $N!$ possible permutations of countries in the value chain, and dynamic programming techniques are ineffective in reducing the dimensionality of those problems.[75]

5.2.2 Horizontal and Export-Platform FDI

Although models combining global production strategies, increasing returns to scale, and trade costs are hard to work with, there is a specific version of those models which has been extensively studied in the literature. This corresponds to the a variant of the models studied above in which $N = 1$, so only final goods are produced, and this is done with local factors of production. Unlike in models of exporting, however, lead firms are not constrained from producing only in the origin country (e.g., the country were they paid the fixed cost of entry). They can instead set up foreign assembly plants to service foreign consumers at a lower marginal cost. These strategies clearly connect with the

---

[75] There is however a key difference between the problem in (52) and the TSP: due to the compounding effect of trade costs, the optimal assignment will put a larger weight on reducing trade costs at relatively downstream stages than at stages further upstream, a result reminiscent to the one in Antràs and de Gortari (2020).

voluminous literature on horizontal FDI and export-platform FDI, which was overviewed in Chapter 2 of the 4th volume of this *Handbook* (see Antràs and Yeaple, 2014, in particular, Section 6.1). Because some of the papers in this literature neatly connect with many firm-level models of GVCs reviewed above, it is however worth briefly reviewing some of their key contributions and insights.

Following the lead of Tintelnot (2017), we envision a two-stage problem in which a lead firm based in country i first activates a set of locations $\mathcal{J}_i \subseteq \mathcal{J}$ after incurring a fixed cost equal to $w_i f_{ik}$ when activating country k, and then decides from which assembly plant in $\mathcal{J}_i$ to sell to consumers in all potential destinations j. The second stage is well captured by the more general problem in equation (51) when setting $N = 1$, which reduces to

$$\ell^j = \arg \min_{\ell^j(N) \in \mathcal{J}_i} p_j^F \left(\ell^j\right) = \arg \min_{\ell \in \mathcal{J}_i} \left\{\tau_{\ell j} a_\ell c_\ell\right\}. \tag{53}$$

when $N = 1$ (where we have dropped n subscripts for simplicity). In words, consumers from j will be serviced from a plant located in the country $\ell^j(N) \in \mathcal{J}_i$ that minimizes the delivery cost of the good $\tau_{\ell j} a_\ell c_\ell$. In order to discern how many assembly plants the firm should set up and where they will be located, one needs a structure to transition from the problem in to a profit function. As in many of the papers reviewed above, Tintelnot (2017) assumes that the firm faces an isoelastic demand function, and the productivities $1/a_\ell$ are drawn from a Fréchet distribution (that is independent across locations). The main new trick that Tintelnot (2017) develops is to assume that firms produce a continuum of consumer goods, which together with the Fréchet assumption, delivers gravity-style equations representing the bilateral sales of all the firm's plants. More specifically, the share of the firm's sales in market j originating from assembly plants in country k is given by:

$$\mu_{ikj} = \frac{T_k \left(\tau_{kj} c_k\right)^{-\theta}}{\sum\limits_{k' \in \mathcal{K}_i} T_{k'} \left(\tau_{k'j} c_{k'}\right)^{-\theta}}, \tag{54}$$

where $\mathcal{K}_i$ is the set of countries or locations where the firm has assembly plants, and where the other parameters are as defined in previous models. Furthermore, one can express the firm's profits conditional on an assembly strategy $\mathcal{K}_i$ as

$$\pi_i\left(\mathcal{K}_i\right) = \kappa \sum_{j \in J} B_j \left(\sum_{k \in \mathcal{K}_i} T_k \left(\tau_{kj} c_k\right)^{-\theta}\right)^{(\sigma-1)/\theta} - w_i \sum_{k \in \mathcal{K}_i} f_{ik}, \tag{55}$$

where κ is a constant, and B_i is a residual demand term. For the price index associated with the bundle of varieties produced by the firm to be bounded, Tintelnot (2017) shows that one needs to impose $\sigma - 1 < \theta$. This parametric restriction implies that the model features 'market cannibalization' effects: a firm may find it optimal to set up a plant in country k to reduce the costs of selling goods to consumers in country k and nearby countries, but such a decision necessarily reduces the marginal benefit of setting up plants in other countries $k' \neq k$. Using the terminology in Antràs et al. (2017), the firm's assembly strategy features substitutability (or decreasing differences) in the entry decisions in alternative markets. Thus, although choosing an assembly strategy amounts to choosing a set among 2^J possible sets, the problem can be in principle be solved using the algorithm suggested in Arkolakis and Eckert (2017). In his empirical application, which focused on the horizontal and export-platform strategies of German multinationals, Tintelnot (2017) instead restricted the analysis to a case in which $J = 12$, so he could solve the problem by brute force.

In recent work, Antràs et al. (2020) develop a multi-country model in which firms choose not only the locations of their various assembly plants, as in the horizontal FDI and export-platform literature, but also the countries from which all those plants import inputs, as in the global sourcing literature. The model in Antràs et al. (2020) constitutes a marriage of the Tintelnot (2017) model of export-platform FDI and the global sourcing framework in Antràs et al. (2017). Their framework delivers simple gravity-style formulas for both firm-level bilateral shipments of consumer goods from any country where a firm assembles finished goods to all other countries in the world, as well as firm-level bilateral purchases of intermediate inputs from countries in a firm's sourcing strategy to each country in which that same firm assembles final goods. Crucially, their framework identifies a natural complementarity between these two decisions – as hinted in prior work by Yeaple (2003) and Grossman et al. (2006) – and thus delivers novel implications for the role of geography in shaping the global production strategies of firms. Intuitively, a richer sourcing strategy reduces marginal costs, increases optimal firm scale, and thus makes a richer assembly strategy more appealing (or its associated fixed costs easier to amortize). Similarly, a richer assembly strategy increases overall firm sales and thus makes a more expansive sourcing strategy more appealing (or its associated fixed costs easier to amortize). Empirically, Antràs et al. (2020) merge US Census domestic and trade data with the US Bureau of Economic Analysis (BEA) comprehensive surveys on multinational activity to document a series of novel facts regarding the global assembly and global sourcing strategies of US-based firms, and they develop new tools to estimate their model structurally and to perform counterfactual exercises that illustrate the rich implications of changes in trade costs on global production patterns. More specifically, due to the coexistence of sources of substitutability (market cannibalization) and complementarity in the model, the problem of determining of a firm's extensive margin decisions does not feature the type of "single-crossing" properties that typically rationalize the use of iterative algorithms to reduce the dimensionality of the problem, as in Jia (2008), Antràs et al. (2017) or Arkolakis and Eckert (2017). To make progress on this issue and render feasible a structural estimation of our model, Antràs et al. (2020) develop a probabilistic approach to solve the firm's extensive margins of global sourcing and global assembly, which smooths out the firm's problem and allows them to characterize its solution by studying and computationally approximating the first-order conditions of this problem via Monte Carlo integration.

We have thus far described recent contributions that tightly connect with other theoretical work described in this Chapter, but it is worth closing this section with a brief description of other recent work on the horizontal or export-platform FDI dimensions of the GVC strategies of lead firms.

First, and as already described in Antràs and Yeaple (2014), Arkolakis et al. (2018) develop a multi-country model in which lead firms decide which country i to enter in, and from which country

k to service consumers in each country j. As in the work of Tintelnot (2017), this appears to be a complex combinatorial problem. Arkolakis et al. (2018) achieve tractability by abstracting away from fixed costs of setting up assembly plants, and only modelling fixed costs of marketing goods in country j regardless of the origin of production k. This basically turns the problem of figuring out the source of goods k as one that minimizes marginal costs, as in the work of Eaton and Kortum (2002). Yet the presence of marketing costs implies that individual firms may only produce in and sell to a subset of countries. By making suitable assumptions about the distribution of productivity across goods and countries, Arkolakis et al. (2018) further show that their framework delivers simple expressions for bilateral trade flows across countries, some of which reflect standard exporting, while the rest reflect export-platform sales.

Arkolakis et al. (2018)'s framework also incorporates iceberg-style costs associated with firms assembling goods in countries other than their country of incorporation. This added cost can be interpreted as a reduced-form way to capture the costs of importing inputs (Ramondo and Rodríguez-Clare, 2013) or knowledge (Keller and Yeaple, 2013) from the headquarters, or perhaps the costs of adapting production to a foreign and unfamiliar environment. An alternative cost of multinational activity features prominently in the recent work of Head and Mayer (2019), who argue that firms also incur adaptation costs when marketing goods in countries distinct from their origin country, regardless of where those goods are produced. They argue that this is an especially important feature of the car industry, and they estimate an industry equilibrium model in which car makers decide on the optimal sourcing of their car models taking into account where their headquarters are located, where they have assembly plants (which are fixed in their model), and where consumers are. Another novel feature of their framework is the inclusion of external economies of scale, which they argue are also a key feature of the industry under study.

5.2.3 Taking Stock

In sum, a growing literature is developing tools to better capture the complex operational decisions of large multinational firms organizing GVCs. Progress in this branch of the literature has been hampered by data availability and by computational complexity.

On the data front, testing models of the decisions of lead firms requires obtaining data on the operations of firms in more than one country, and thus standard data sources used by trade economists, such as customs data or industrial census datasets are not suitable for this goal. Nevertheless, progress has been made by either exploiting survey data on the outward operations of multinational firms, and also by industry-specific studies relying on more granular data for specific sectors, such as the car industry data in Head and Mayer (2019).

On the computational front, and as we have made clear above, lead firm problems face highly complex decisions when designing their supply chains. The economics literature has largely been focused on finding suitable environments in which these decisions can be qualitatively characterized or computationally simplified, but it is hard to envision at this point that this agenda will lead to successful unified quantitative models of the decisions of lead firms. Our sense is that, sooner or later, this literature will need to close the gap with the parallel literature in supply chain management in the Operations Research field, which has long adopted heuristic methods to guide the optimal design of supply chains (see Vidal and Goetschalckx, 1997, for a review).

5.3 Organizing Relational GVCs

Although the research reviewed in the last few sections has provided valuable novel insights regarding the emergence and implications of GVCs, modeling global production sharing as simply an increase in the extent to which intermediate inputs are exported or imported (or to which multinationals set up foreign export platforms) misses important distinctive characteristics of the recent rise of GVCs. Three of these distinctive features are particularly important.

First, finding suitable suppliers of parts and components or suitable buyers of one's products is costly, or in economic lingo, there are search frictions. As a result, the fixed costs of exporting and importing we have been referring to in previous sections, are better understood as sunk costs, which naturally create a "stickiness" among participants in a GVC.

Second, GVC participants often undertake numerous relationship-specific investments (such as purchasing specialized equipment or customizing products) which would obtain a much-depressed return were GVC links to be broken. The need to customize inputs adds to search frictions in creating a 'lock-in' effect that further contributes to tie together the different agents in a GVC.

Third, the prevalence of lock-in effects within GVCs is made particularly relevant by the limited contractual security governing transactions within these chains. There are in turn two reasons why GVC participants perceive contractual insecurity. On the one hand, GVCs often involve transactions for which a strong legal environment is particularly important to bind producers together and to preclude technological leakage. On the other hand, GVCs often flow into countries with weak contracting institutions that do not offer the same contractual safeguards that typically accompany similar exchanges occurring in rich countries. As a result, GVC participants are often left to employ repeated interactions among them to build a governance that provides implicit contract enforcement. As in the case of matching frictions and relationship-specificity, this force contributes to the "stickiness" of GVC links. In sum, GVC *relationships* matter, and thus this branch of this literature has come to be referred to as studying *relational GVCs*.

In this section, we will briefly overview theoretical work that has attempted to shed light on the workings of relational GVCs. We will first study the role of relationship-specificity and contractual frictions in shaping the location and firm-boundary decisions in GVCs. We will in turn do so in two steps. First, we will follow Chor and Ma (2020) and consider a spider-like model of backward GVC participation à la Antràs et al. (2017) expanded to include contractual frictions and firm boundary decisions. Next, we will build on Antràs and Chor (2013) and Alfaro et al. (2019) and perform an analogous exercise for snake-like sequential production processes. We will close this section with a succinct account of additional work highlighting the role of search frictions and of relational contracting in shaping GVCs.

Before diving in, we should stress that we focus on recent developments in this branch of the literature. For earlier work on contractual frictions and firm boundary choices in international trade contexts, we refer readers to Chapter 2 of Volume 4 of this *Handbook* (see Antràs and Yeaple, 2014, in particular, Section 7).

5.3.1 Contractual Frictions and Firm Boundaries in Spiders

Let us return to the model of backward GVC participation developed in section 5.1.2. Remember that the framework features a final-good sector with CES preferences over differentiated varieties and an intermediate-input sector that provides differentiated input varieties to the final-good sector, which combines them with labor in production according to equation (38). The bundle of intermediate inputs is also characterized by a CES aggregator, as in equation (39). The final good-sector features increasing returns to scale technologies and is monopolistically competitive, while the upstream sector produces under constant returns to scale with productivity levels drawn from a Fréchet distribution. Final-good varieties are nontradable, but intermediate inputs can be traded across borders with associated iceberg trade costs τ_{ij}^{I}.

Chor and Ma (2020) embeds a property-rights model of firm boundaries à la Antràs (2003) and Antràs and Helpman (2004) into this framework.[76] More specifically, the following new assumptions are made. First, each input variety ϖ is produced combining headquarter services and a manufacturing input provided by the supplier according to a Cobb-Douglas technology:

$$y_{ij}^{s}\left(\omega,\varpi\right)=z_{ij}^{I}\left(\omega,\varpi\right)\left(h_{ij}\left(\omega,\varpi\right)\right)^{\eta}\left(m_{ij}\left(\omega,\varpi\right)\right)^{1-\eta}, \tag{56}$$

where η reflects the headquarter-intensity of input production. Although $h_{ij}\left(\omega,\varpi\right)$ is provided by the final-good producer located in i, Chor and Ma (2020) assume, following Antràs (2003), that $h_{ij}\left(\omega,\varpi\right)$ is produced in country j using factors of production in that country. Second, both headquarter services and the supplier input are relationship-specific in the sense that they are each customized as inputs for the final-good producers' consumption variety. Third, certain aspects of the production of both headquarter services and of input manufacturing cannot be specified in a fully enforceable manner in an initial contract between the final-good producer and the supplier. A simple way to model this, following Acemoglu et al. (2007) and Antràs and Helpman (2008), is to assume that only a fraction μ_{ij}^{h} of the tasks that go into producing headquarter services and a fraction μ_{ij}^{m} of the tasks that go into producing manufacturing inputs are contractible. With a symmetric Cobb-Douglas technology across tasks, this amounts to rewriting technology in (56) as

$$y_{i}^{s}(\varpi)=z_{ij}^{I}\left(\omega,\varpi\right)\left(\left(h_{ij}^{c}\left(\omega,\varpi\right)\right)^{\mu_{ij}^{h}}\left(h_{ij}^{n}\left(\omega,\varpi\right)\right)^{1-\mu_{ij}^{h}}\right)^{\eta}\left(\left(m_{ij}^{c}\left(\omega,\varpi\right)\right)^{\mu_{ij}^{m}}\left(m_{ij}^{n}\left(\omega,\varpi\right)\right)^{1-\mu_{ij}^{m}}\right)^{1-\eta},$$

where $h_{ij}^{c}\left(\omega,\varpi\right)$ and $m_{ij}^{c}\left(\omega,\varpi\right)$ are the symmetric investments in contractible tasks, and $h_{ij}^{c}\left(\omega,\varpi\right)$ and $m_{ij}^{c}\left(\omega,\varpi\right)$ are the analogous investments in non-contractible tasks. Finally, because some investments are not contractible ex-ante, one needs to specify how the terms of exchange will be

---

[76] To be precise, Chor and Ma (2020)'s framework incorporates multiple upstream sectors, but we only model one here for simplicity.

determined ex-post, once all investments have been incurred.[77] As is standard in the literature, Chor and Ma (2020) characterize this ex-post bargaining using the Nash Bargaining solution and assume symmetric information between headquarters and the various suppliers. In that bargaining, the final-good producers walks away with a share β_{ij} of the surplus from the relationship, with this surplus in turn related to the contribution of all the other suppliers into production. The share β_{ij} may be shaped by primitive bargaining power or relationship specificity asymmetries (see Antràs, 2016; Eppinger and Kukharskyy, 2020), but crucially and following the property-rights approach, it is also shaped by firm boundary decisions. When the supplier is integrated, the final-good producer obtains a share β_{ij}^{V} of surplus that is higher than the share β_{ij}^{O} it obtains when the supplier is a stand-alone firm.

The Chor and Ma (2020) model is much richer than the underlying Antràs et al. (2017) framework, but it is simpler in an important sense: Chor and Ma (2020) abstract from fixed costs of importing, and thus firms source inputs from all countries in the world. Nevertheless, firms' sourcing strategies are richer in the sense that the firm has $2J$ potential sources for each input, corresponding to the J countries and two organizational forms (vertical integration versus outsourcing). To capture the intuitive notion that the productivity of integrated and independent suppliers in a given country j should be correlated, Chor and Ma (2020) assume that the productivity term $z_{ij}^{I}\left(\omega,\varpi\right)$ in (56) is drawn independently for each ω and ϖ from a "nested-Fréchet" distribution with cumulative distribution function (cdf):

$$\begin{aligned}\Pr\left(z_{1j}^{V}\left(\omega,\varpi\right)\leq z_{1j}^{V},z_{1j}^{O}\left(\omega,\varpi\right)\leq z_{1j}^{O},\ldots,Z_{JjO}^{O}\leq z_{JjO}^{O}\right)\\ =\exp\left\{-\sum_{i=1}^{J}T_i\left(\left(z_{ij}^{V}\right)^{-\frac{\theta}{1-\lambda_i}}+\left(z_{ij}^{O}\right)^{-\frac{\theta}{1-\lambda_i}}\right)^{1-\lambda_i}\right\},\end{aligned}$$

where $T_i^k > 0$, $\theta^k > 1$ and $0 < \lambda_i < 1$ for each source country i. The parameters λ_i govern the correlation in the productivity draws obtained by stand-alone and integrated suppliers, with $\lambda_i = 1$ implying an identical productivity, and $\lambda_i = 0$ for all countries i corresponding to the special case where the $2J$ draws are each from independent Fréchet distributions with cdf: $\exp\left\{-T_i(z_{ij})^{-\theta}\right\}$. This specification delivers a closed-form expression for sourcing shares that has an intuitive nested logit form: The share of inputs obtained from country i under (say) integration is equal to the share sourced from country i, multiplied by the share sourced under integration conditional on having chosen country i. Furthermore, these shares are not only shaped by standard parameters, such as levels of technology, trade costs and wages, but also by institutional or contractual parameters, such as the degrees of contractibility μ_{ij}^{h} and μ_{ij}^{m}, and the bargaining parameters β_{ij}^{V} and β_{ij}^{O}.

The fact that Chor and Ma (2020) ignore the extensive margin of which source countries and organizations to activate allows them to neatly characterize the general equilibrium of the model, and compare it to recent quantitative models in the field. For instance, the framework delivers an

---

[77] The initial contract specifies binding investment levels for all contractible tasks, as well as a lump-sum transfer between the agents.

expression for the welfare gains from trade that is akin (in the limit case where all inputs are fully contractible) to that in Arkolakis et al. (2012). They also show that their framework is amenable to the use of the hat-algebra approach to counterfactuals in Dekle et al. (2008) and Caliendo and Parro (2015), but crucially, their framework suggests that whether trade shares take the form of intrafirm or arm's-length imports is consequential for welfare. Chor and Ma (2020) are then able to perform counterfactuals that evaluate the welfare consequences of improving the contractual environment, as well as studying the way in which the magnitude of the gains from trade interacts with the level of contracting institutions.

Although we have highlighted the work of Chor and Ma (2020), it is worth closing this section outlining recent work that has similarly explored how contractual frictions shape the sourcing decisions of firms, and how those decisions in turn shape the consequences of trade integration or of changes in the contractual environment. Many of the early contributions to this literature – which focused on low-dimensional models – are overviewed in Antràs and Yeaple (2014) and Antràs (2016). The work of Boehm (2020) and Boehm and Oberfield (2020) is much closer in spirit to the work of Chor and Ma (2020) in that they also provide welfare assessments of contracting frictions, with Boehm (2020) in particular doing so in a full multi-country, multi-sector general equilibrium setting. That said, contracting frictions in these papers are modeled based on transactions-cost theory, in situations in which the firm-supplier relationship features a one-sided (rather than bilateral) holdup problem.

5.3.2 Contractual Frictions and Firm Boundaries in Snakes

We next turn to a parallel set of studies of how contractual frictions shape the location and organization of GVCs, but this time we focus on purely sequential production process. We begin overviewing the work of Antràs and Chor (2013) and Alfaro et al. (2019) who develop and test the implications of a property-rights model of sequential production.

The setting is similar to the models described in Sections 4.2 and 5.2.1, except that production stages are characterized as a continuum. More specifically, Antràs and Chor (2013) focus on the problem of a final-good producer facing an isoelastic demand for its product, that is seeking to optimally organize a sequential manufacturing process that requires the completion of a unit measure of production stages. These stages are indexed by $i \in [0,1]$, with a larger i corresponding to stages further downstream and thus closer to the finished product. Denote by $x(i)$ the value of the services of intermediate inputs that the supplier of stage i delivers to the firm. Final-good production is then given by:

$$y^F(\omega) = z_i^F(\omega) \left(\int_0^1 y^s(\omega, \varpi)^\rho \mathcal{I}(\varpi)\, d\varpi \right)^{1/\rho}, \tag{57}$$

where $z_i^F(\omega)$ is a productivity parameter, $\rho \in (0,1)$ is a parameter that captures the (symmetric) degree of substitutability among the stage inputs (as in equation (38)), and $I(\varpi)$ is an indicator function that takes a value of 1 if input ϖ is produced after all inputs $\varpi' < \varpi$ have been produced, and a value of 0 otherwise. It is this last indicator function $\mathcal{I}(\varpi)$ that makes the production technology inherently sequential.

The contractual aspects of the model are in many ways analogous to those discussed above in the Chor and Ma (2020) framework. The different stage inputs are provided by suppliers, who each undertake relationship-specific investments to make their components compatible with those of other suppliers along the value chain. The setting is one of incomplete contracting, in the sense that contracts contingent on whether components are compatible or not cannot be enforced by third parties. As a result, the division of surplus between the firm and each supplier is governed by bargaining, after a stage has been completed and the firm has had a chance to inspect the input. At that point, the firm and the supplier negotiate over the division of the incremental contribution to total revenue generated by supplier i, independently from the bilateral negotiations that take place at other stages (see Antràs and Chor, 2013, for alternative formulations of the bargaining protocol). In the initial stage of the model, the firm must decide which input suppliers (if any) to own along the value chain. As in the property-rights theory, the integration of suppliers does not change the space of contracts available to the firm and its suppliers, but it affects the relative ex-post bargaining power of these agents. Vertical integration confers the final-good producer higher bargaining power than outsourcing.

In order to solve for the subgame perfect equilibrium of the above game, Antràs and Chor (2013) note that the quasi-rents over which the firm and the supplier at position ϖ in the value chain negotiate are given by the incremental contribution to total revenue generated by supplier ϖ at that stage, which in turn are given by:

$$r'(\omega,\varpi) = \kappa \left(z_i^F(\omega) \right)^{\rho} \left(r(\omega,\varpi) \right)^{1-\frac{\sigma-1}{\sigma\rho}} \left(y^s(\omega,\varpi) \right)^{\rho}, \tag{58}$$

where $r(\omega,\varpi)$ is the revenue *secured* by the final-good producer up to stage ϖ. As highlighted by Antràs and Chor (2013), whenever $\sigma > 1/(1-\rho)$, the investment choices of suppliers are *sequential complements* in the sense that higher investment levels by prior suppliers increase the marginal return of supplier ϖ's own investment $y^s(\omega,\varpi)$. Conversely, if $\sigma < 1/(1-\rho)$, investment choices are *sequential substitutes* because high values of upstream investments reduce the marginal return to investing in $y^s(\omega,\varpi)$. Because the supplier at position ϖ chooses $y^s(\omega,\varpi)$ to maximize $(1-\beta(\omega,\varpi))\, r'(\omega,\varpi) - c(\omega,\varpi) y^s(\omega,\varpi)$, where $c(\omega,\varpi)$ is the marginal cost of investment, equation (58) illustrates the trickle-down effect that upstream investment inefficiencies can have on downstream stages.

Exploiting the recursive structure of the model, Antràs and Chor (2013) characterize the optimal division of surplus along the chain. The key result in their paper is that the relative size of the input and final-good elasticities of substitution, respectively $\sigma_\rho = 1/(1-\rho)$ and σ, governs whether the incentive for the final-good producer to retain a larger surplus share increases or decreases along the value chain. Intuitively, when σ is high relative to σ_ρ, investments are sequential complements, and high upstream values of $\beta(\omega,\varpi)$ are particularly costly since they reduce the incentives to invest not only of these early suppliers but also of all suppliers downstream. Conversely, when σ is small relative to σ_ρ, investments are sequential substitutes, and low values of $\beta(\omega,\varpi)$ in upstream stages are now relatively detrimental, since they reduce the incentives to invest for downstream suppliers, who are already underinvesting to begin with.

Alfaro et al. (2019) develop several extensions of the Antràs and Chor (2013) model that are relevant for their firm-level empirical analysis. First, they introduce asymmetries across inputs and map them to variation across inputs in the degree of contractibility. Second, they incorporate heterogeneity across final good producers in their core productivity, while introducing fixed costs of integrating suppliers, as in Antràs and Helpman (2004). They then show how such productivity differences influence the number of stages that are integrated, and hence the propensity of the firm to integrate upstream relative to downstream stages. Finally, they consider a scenario in which integration is infeasible for certain segments of the value chain, for example, due to exogenous technological or regulatory factors, and demonstrate that the model's predictions continue to describe firm boundary choices for those inputs over which integration is feasible.

Although Antràs and Chor (2013) and Alfaro et al. (2019) abstract from the study of location choices, their results have potentially interesting implications for the choice between domestic and foreign sourcing whenever these sourcing strategies are associated with different levels of contract enforcement. To see this, consider the case in which contracting in domestic transactions is complete, while foreign sourcing is associated with incomplete contracting (as in Antràs, 2005). The results in Antràs and Chor (2013) then suggest that, in the sequential complements case ($\sigma > \sigma_\rho$), foreign sourcing is particularly unappealing in upstream stages. Thus, if domestic and foreign sourcing coexist along the value chain, then only relatively downstream inputs will be offshored. Conversely, in the sequential substitutes case, ($\sigma < \sigma_\rho$) one would expect relatively upstream stages to be offshored. In sum, the model predicts that the 'upstreamness' of an input should be a relevant determinant of the extent to which it is procured from foreign suppliers, with the sign of that dependence being crucially shaped by the relative size of σ and σ_ρ.

In largely contemporaneous work, Fally and Hillberry (2018) developed an alternative framework illustrating the consequences of contractual frictions for the location and organization of GVCs. Their framework in turn builds on the insightful transaction-cost model in Kikuchi et al. (2018). In that framework, production of a final good requires (again) that a continuum of stages or tasks be executed in a pre-determined order. A set of identical firms can produce any set of tasks with decreasing returns with respect to the measure of tasks produced, which fosters specialization across firms. To put a check on specialization and generate firms that produce a measurable set of tasks, Kikuchi et al. (2018) assume that firm-to-firm transactions involve a cost that is proportional to the value and price of the good at the time of delivery. Kikuchi et al. (2018) provide a sharp characterization of this problem, in part using recursive methods. For reasons analogous to those in Costinot et al. (2013) and Antràs and de Gortari (2020), the resulting allocation has relatively small firms in the upstream stages of production, with firm size growing monotonically as one moves to more and more downstream stages. The authors claim that the model can match the size distribution of firms and they further provide a set of comparative statics with respect to transaction costs, and parameters of the cost function. Fally and Hillberry (2018) extend this setup to an international setting with costly trade, and develop implications for within-chain comparative advantage. Fally and Hillberry (2018) demonstrate that their framework delivers a positive relationship between a country-industry pair's upstreamness measure and its gross-output-to-value-added ratio, and they provide empirical evidence consistent with it.

5.3.3 Search Frictions

Beyond frictions associated with incomplete contracting, the literature on the international organization of production has also stressed the role of search frictions in shaping the emergence and sustainability of GVC links. The standard way to model these frictions is by assuming that the type of fixed-cost investments firms incur to match with GVC partners, as modelled in Sections 5.1.1, 5.1.2 and 5.1.3, only deliver matches with a probability governed by the relative mass of firms searching for matches in both sides of the market. The work of Grossman and Helpman (2005) constitute an early contribution to this literature in a simple two-country general equilibrium model, while in more recent work, Eaton et al. (2014), Allen (2014), Krolikowski and McCallum (2018) and Lenoir et al. (2019) adopt a similar approach in more complex, multi-country quantifiable models.[78]

The introduction of search frictions enriches the set of predictions emanating from models of firm-level GVC participation. Without delving into the technical details of these models, we would highlight four main new sets of insights. First, other things equal, it is clear that search frictions reduce the attractiveness of engaging in GVC activity, and might lead some firms to opt out of it when they would have found it profitable to participate in the absence of these frictions. This in turn carries consequences for the welfare responses to trade shocks and for the trade elasticity, relative to models without search frictions (see Krolikowski and McCallum, 2018). Second, in the presence of increasing returns in the matching function, this line of models can generate multiple equilibria and *waves* of GVC participation, as the entry of some firms may generate a positive spillovers on the entry of other firms (see McLaren, 2000; Grossman and Helpman, 2002). Third, because the fixed costs associated with matching with other producers are sunk in nature, this line of models also tends to feature hysteresis in the margins of trade (see Eaton et al., 2014), which in turn has implications for how the geography of GVCs responds to shocks, such as the current COVID-19 pandemic (see Antràs, 2020). Fourth, this hysteresis can also be interpreted as a form of lock-in effect, which binds buyer-seller pairs together, and thus aggravates the type of contractual frictions outlined in Sections 5.3.1 and 5.3.2.

A natural way to reduce search frictions in finding suitable GVC partners is to rely on specialized intermediaries. It is thus not surprising that recent work on intermediation in international trade has also developed frameworks in which search frictions are prominent, as in the work of Antràs and Costinot (2011), Dasgupta and Mondria (2018), or Startz (2016).

---

[78] See also McLaren (2000) and Grossman and Helpman (2002) for even earlier contributions with search frictions. Other authors refer to the costs of matching as search costs – see Monarch (2020) or Antràs et al. (2020)– but in this section we focus attention on settings in which the probability of a match is a function of the sets of agents searching.

5.3.4 Relational Contracting

As we have argued in Sections 5.3.1 and 5.3.2, the relational nature of GVCs highlights the role of institutional quality as a significant determinant of GVC participation. Nevertheless, the same forces that make relational GVCs rely intensively on institutional quality, such as the lock-in effects created by relationship-specific investments and by search frictions, also make GVC links particularly "sticky", which fosters the emergence of reputational mechanisms of cooperation which might partly substitute for the absence of formal contracting.

An extreme version of this type of relational contracting arises when parties involved in a GVC altogether by-pass the market mechanism and decide to transact within firm boundaries, as in the work outlined in Sections 5.3.1 and 5.3.2. Nevertheless, the internalization of transactions in a GVC is just one of the many organizational responses to the contractual vagaries associated with cross-border transactions. In an influential study in the management literature, Gereffi et al. (2005) elaborate on a much more extensive taxonomy of potential governance forms within GVCs, and various researchers have built on their work to shed light on the relative prevalence of these governance forms through a number of interesting case studies (see Van Biesebroeck and Schmitt, 2020, for a recent example in the economics literature).

Because trade economists typically favor modes of governance that can be identified in the data across various industries, the literature in international trade has largely focus on exploring the emergence and consequences of relational contracting. This is perceived as an intermediate option between vertically integrated GVC links and spot market transactions with suppliers. We reviewed the burgeoning empirical trade literature on relational contracting in Section 3.3. Here, we simply outline some of the theoretical insights from that literature.

It is useful to begin by identify two broad approaches to modeling how relational contracting shapes GVC participation and trade flows. The first approach, which one might call the 'adverse selection' approach, considers environments in which certain GVC participants – say buyers – come in two fixed types: honest and dishonest. Honest buyers always honor contracts (say paying for the delivered goods), even when the contractual environment is weak and cannot always impose penalties on misbehavior. On the other hand, dishonest agents misbehave when given a chance, which in these models occurs with certain probability. In this environment, repeated contracting allows sellers to better learn and identify whether buyers are honest or dishonest. More precisely, starting from a prior, sellers update their belief of the buyer being honest as long as no misbehavior is observed in equilibrium. Yet, when misbehavior is observed, sellers immediate infer they are dealing with a dishonest buyer, and optimally discontinue the relationship. This line of models is developed in Antràs and Foley (2015) and Araujo et al. (2016), and it has been further developed and taken to the data by Monarch and Schmidt-Eisenlohr (2017), a paper we discussed in some detail in section 3.3. A distinguishing feature of these line of models are that they naturally generate an increasing volume of trade as relationship age increases, with the rate of growth in firm-to-firm trade being larger, the weaker the contractual environment.

A second line of models, which we can refer to as 'moral hazard' approaches are more in line with the models of contractual frictions reviewed in sections 5.3.1 and 5.3.2. These frameworks build on the relational contracting literature (see MacLeod and Malcomson, 1989; Levin, 2003; Board, 2011), in which agents undertake noncontractible investments, and in which repeated interactions may allow them to sustain cooperation under the threat of reversion to a non-cooperative equilibrium. This literature is largely concerned with characterizing the range of parameter values for which cooperation can be sustained. A particularly noteworthy contribution is the work of Defever et al. (2016) who study a model of global sourcing (or backward GVC participation) very much in the spirit of the double-sided hold-up problem in Antràs (2003), Antràs and Helpman (2004) and Chor and Ma (2020) (see Section 5.3.1). An intuitive insight from these models is that relational contracting is more likely to be used whenever agents are relatively patient, in which case the costs of reverting to a non-cooperative equilibrium are higher. Less trivially, these frameworks also illustrate how weak contracting, by depressing the payoffs under non-cooperation, renders relational cooperation more beneficial. In this sense, formal and informal contracting appear to be substitutes. An unappealing (or counterfactual) implication of these line of frameworks – at least in their most stripped-down form – is that they tend to general first-best investment levels and trade volumes from the onset of a relationship. To remedy this, some authors such as Macchiavello and Morjaria (2015), have considered environments that blend the adverse selection and moral hazard approaches.[79]

In closing, it is also worth mentioning some theoretical work, building on Baker et al. (2002), that has developed frameworks in which firms not only choose between engaging in spot versus relational contracting, but also consider the possibility of internalizing transactions, as in the work reviewed in Sections 5.3.1 and 5.3.2. Notable contributions to this literature include the work of Kukharskyy (2016) and Kamal and Tang (2014).

6 Trade Policy in the Age of GVCs

We would argue that, relative to the work on the measurement and modeling of GVCs, our profession's understanding of the policy implications of the rise of GVCs is much less fully developed. How does the use of traditional instruments of trade policy – tariffs, quantitative restrictions, or regulatory standards – affect the volume of trade and social welfare in a world of GVCs relative to a world where trade is exclusively in final goods? How does the rise of GVCs affect our understanding of what constitutes optimal trade policy? Answers to these questions are particularly relevant in current times, when trade policy discussions are as salient as they have been in the last fifty years.

We will structure the discussion along the following main themes. First, in sections 6.1, 6.2, and 6.3, we review work studying the implications and optimal design of trade policy in competitive environments featuring final-good trade as well as intermediate-input trade. Second, in Section 6.4, we study the role of political-economy forces in shaping the structure of protection in upstream and downstream markets. Third, we extend the analysis to richer frameworks, such as frameworks with differentiated domestic and foreign value added in production (section 6.5), general-equilibrium

---

[79] See Gil (2011) for a study of the interplay between formal and informal contracting in the context of the movie industry.

Ricardian models featuring product differentiation in upstream and downstream markets (section 6.6), models featuring imperfect competition (section 6.7), and models of relational GVCs (section 6.8).

Although we will connect at times with selected empirical papers, the focus of this section is admittedly theoretical in nature. In recent years, a number of interesting empirical papers have been developed to study the implications of trade protection for the geography of GVCs. Conconi et al. (2018) use rich data to study the implications of NAFTA's "Rules of Origin" for intermediate input trade, and show that these RoO led to a sizable decrease in imports of intermediate goods into Mexico from third countries relative to imports from the U.S. and Canada. Relatedly, Vandenbussche and Viegelahn (2018) document that following the imposition of antidumping duties, Indian firms reduce their use of protected inputs on average by 25–40%, relative to other inputs. Bown et al. (2020) study the impact of U.S. antidumping duties against Chinese imports during the period 1988-2016 (and also during the recent U.S.-China trade war), and find that they had a significantly negative impact on employment in downstream industries, with no counterbalancing positive effect in protected industries. Similar results were obtained by Barattieri and Cacciatore (2020) for the period 1994-2015, and by Flaaen and Pierce (2019) when focusing on the Trump tariffs (and subsequent foreign retaliatory tariffs) instituted since 2018. Using a shift-share design based on pre-period shares constructed with detailed U.S. firm-level data from 2016, Handley et al. (2020) similarly report a negative association between firm-level export growth and the extent to which a firm's input bundle in 2016 included inputs that eventually faced higher tariffs during the 2018-19 tariff war.[80] These recent studies were bolstered by the fact that the recent Trump tariffs, unlike most previous protectionist episodes, disproportionately affect intermediate inputs. In particular, Bown and Zhang (2019) calculate that approximately 60 percent of the Trump tariffs through 2018 were on inputs, and affected nearly 20 percent of all US imports of intermediate inputs.

6.1 Effective Rate of Protection and Tariff Escalation

As in the rest of this survey, our focus is largely on work carried out (roughly) in the last ten years, but it is worth beginning this section with an overview of some leading themes on the older trade policy literature that are very much related to intermediate input trade and to global value chains.

Consider first the literature on effective rates of protection, which is exemplified by the seminal work of Corden (1966). This literature is concerned with the implications of intermediate input trade for the incidence of import tariffs. More precisely, Corden (1966)'s definition of the effective rate of protection is "*the percentage increase in value added per unit in an economic activity which is made possible by the tariff structure relative to the situation in the absence of tariffs.* (p. 222)" To formally study this concept, consider a simple partial-equilibrium environment in which final output in a given industry is produced with local value added and with a bundle of intermediate inputs, as in many of the models studied in previous sections. It is assumed that local value added

---

[80] Other notable studies of the implications of the tariffs introduced by the Trump administration include the work of Amiti et al. (2019), Fajgelbaum et al. (2020), and Flaaen et al. (2020).

subsumes any usage of local intermediate inputs, so the bundle of intermediate inputs is imported from the rest of the world.[81] Assuming zero profits, we have

$$p_F = va_v + p_I a_I,$$

where p_F is the price of the final good, a_v and a_I are the unit value-added and input-bundle requirements, and v and p_I are the price for value added and the input bundle.

Suppose that the country under study is a small open-economy, so in an untaxed equilibrium, we have that p_F and p_I correspond to the world prices for these inputs. Now, consider the implications of levying ad-valorem import tariffs t_F on the final good, and t_I on the input bundle. Given the small-country assumption, the price of the final good will increase to $p'_F = t_F p_F$, while the price of the input bundle, will rise to $p'_I = t_I p_I$. Naturally, such protection will benefit local value added in that sector, which can see its remuneration rise. But, by how much? Holding a_v and a_I constant, it is straightforward to see that v can increase by

$$\frac{v'}{v} = t_F + (t_F - t_I)\,\frac{p_I a_I}{v a_v}. \tag{59}$$

The effective rate of protection is thus higher than t_F provided that $t_F > t_I$, and is also increasing in the importance of inputs in production. As an example, if imported inputs are untaxed, and the ratio of value added to gross output is 1/2, as roughly observed in the data, the effective rate of protection is *twice* that implied by the tariff on final goods. Furthermore, it is straightforward to extend formula (59) to the case of multiple intermediate inputs indexed by s, facing heterogeneous tariffs levels

$$\frac{v'}{v} = t_F + \sum\nolimits_s (t_F - t_s)\,\frac{p_s a_s}{v a_v}. \tag{60}$$

Around the time Corden was developing these results, several authors put these and related formulas to use and showed that effective rates of protection were indeed much higher than those implied by final-good tariffs. Among others, Balassa (1965) provided evidence for relatively advanced economies (United States, United Kingdom, the European Common Market, Sweden and Japan) in 1962, while Balassa (1971) complemented this evidence with calculations for a few developing economies (Brazil, Chile, Mexico, West Malaysia, Pakistan, the Philippines, and Norway), which resulted in even higher (and, in some cases, remarkably high) effective rates of protection.[82]

As is clear from equations (59) and (60), the wedge between nominal and effective rates of protection is a reflection of the fact that tariffs applied on final goods are typically higher than those applied on intermediate inputs. This phenomenon is often referred to as 'tariff escalation' and it has been documented in a large number studies, from the early work by Travis (1964) and Balassa (1965),

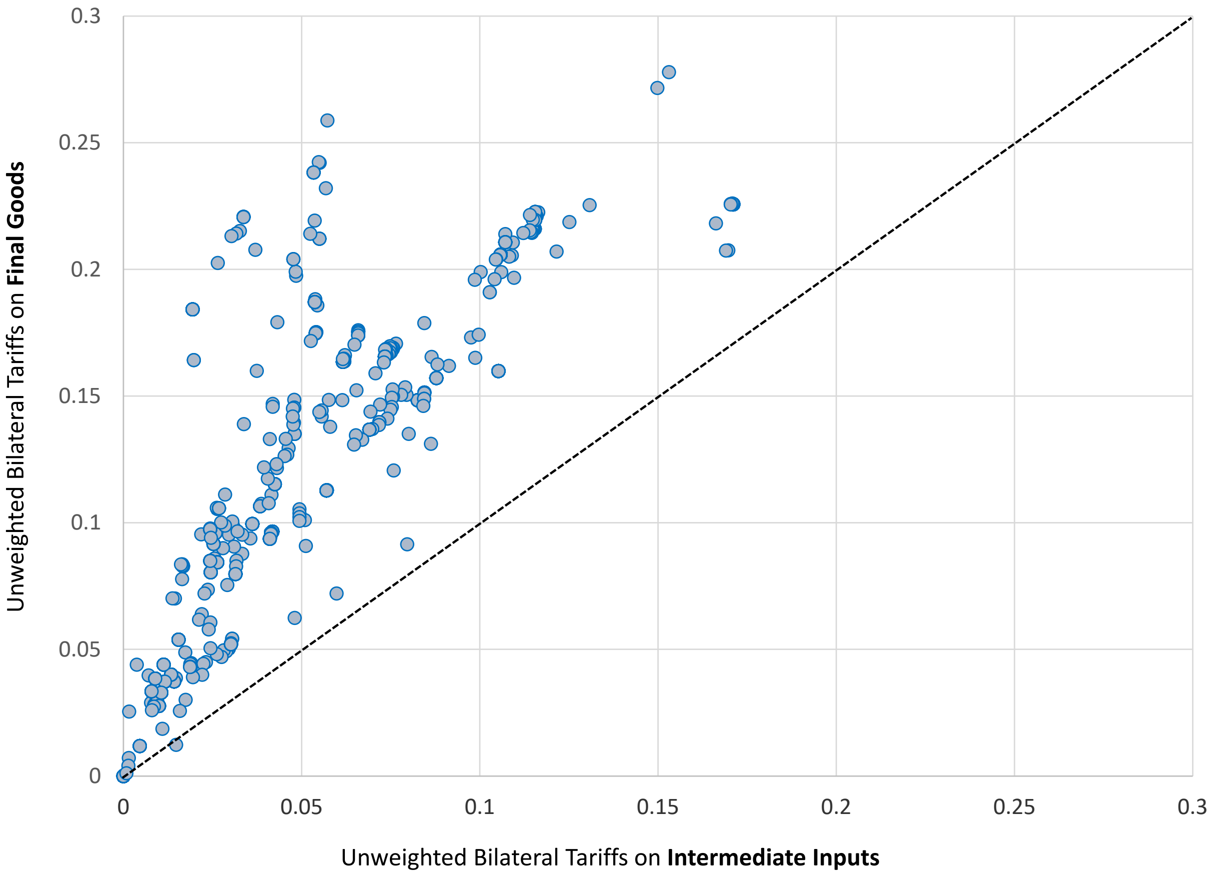
---

[81] This implicitly assumes that local inputs, if any, are produced using *only* local value added, which is inconsistent with roundabout models, such as Caliendo and Parro (2015).

[82] There is also an extensive theoretical literature on the robustness of the simple equations (59) and (60) to alternative environments. See Ethier (1977) for a particularly critical overview, and Anderson (1998) for a lucid rejoinder.

to the more recent calculations by Bown and Crowley (2016) and Shapiro (2020). To illustrate this phenomenon, Figure 7 depicts applied bilateral tariffs on final goods versus intermediate inputs for 37 countries, as computed by Shapiro (2020), where the distinction between a final good and an input is drawn based on the UN BEC end-use classification. As is clear from the figure, all but one of the scatter points are above the 45-degree line, and in many cases by a wide margin. Shapiro (2020) also finds a smoother negative correlation between tariffs and the upstreamness of a sector, as measured by Antràs et al. (2012).

Figure 7: Tariff Escalation



The existence of a clear pattern of tariff escalation explains why effective rates of protection, as measured in equations (59) and (60), appear larger than nominal rates of protection on final goods. This still leaves open the question of what is the policy relevance of this finding. Are high effective rates of protection bad for economic welfare? Is tariff escalation consistent with the tariffs on final goods and on inputs that a social planner would set? The next sections will attempt to provide tentative answers to these questions.

6.2 Baseline: A Simple Roundabout Model

We begin by analyzing optimal trade policy in a partial-equilibrium, constant-returns-to-scale, perfectly competitive roundabout environment with input-output links. The framework builds on those in Cadot et al. (2004) and Gawande et al. (2012) – which themselves extend the classical Grossman and Helpman (1994) 'protection for sale model' – but we relax the assumption of the country under study being a small open economy, and for now we do not model political economy biases in the setting of policies.

We consider a two-country environment with a Home country, the focus of the analysis, and a Foreign one, which sometimes we refer to as the Rest of the World (RoW). The Home country is populated by a continuum of measure 1 of individuals with identical quasi-linear preferences:

$$U(c) = c_0 + \sum_{s=1}^{S} u_s(c_s),$$

where c_0 denotes consumption of the outside numéraire good, which is costlessly traded and not subject to tariffs. Within each 'outside' sector s, Home and Foreign goods are assumed to be perfect substitutes. Provided that income is large enough, we can characterize the demand side of the model via simple sectoral demand functions that are only a function of each sector's price

$$c_s = d_s\left(p_s\right) \qquad \text{for } s = 1, 2, ..., S,$$

given that the outside good eliminates all income effects, and preferences are separable across sectors. Consumer surplus in sector s is in turn given by $S_s\left(p_s\right) = u_s\left(c_s\left(p_s\right)\right) - p_s c_s\left(p_s\right)$. Preferences are analogous in the Foreign country, leading to a similarly downward-sloping demand for goods in sectors $s = 1, 2, ..., S$.

On the supply side, the numéraire good is produced one-to-one with labor, which pins down the wage rate to 1 in all countries. Non-numéraire goods are produced combining labor, sector-specific capital, and the intermediate inputs consisting of output from all other sectors of the economy. To simplify matters, and following Cadot et al. (2004) and Gawande et al. (2012), we assume that inputs are used in fixed proportions, so we can write the rent function $\Pi_s\left(p_s\right)$ for the capital specific to sector s as

$$\Pi_s\left(p_{1,\ldots,}p_S\right) = \left(p_s - \sum\nolimits_{r=1}^{S} a_{rs} p_r\right) x_s\left(k_s, \ell_s\right) - \ell_s, \tag{61}$$

where x_s is output in sector s, a_{rs} is the fixed requirement of units of good r used as inputs in producing good s, k_s is the fixed amount of sector-specific capital in sector s, and ℓ_s is the amount of labor hired in sector s.

Due to the quasi-linearity of preferences, we can write Home welfare associated with a given vector $\mathbf{p} = (p_{1,\ldots}, p_S)$ of domestic prices and a given vector $\mathbf{p}^* = \left(p^*_{1,\ldots}, p^*_S\right)$ of foreign prices as

$$W(\mathbf{p}, \mathbf{p}^*) = 1 + \sum_{s=1}^{S} \Pi_s(\mathbf{p}) + \sum_{s=1}^{S} S_s(p_s) + \sum_{s=1}^{S} (p_s - p^*_s) \left(c_s(p_s) + \sum\nolimits_{r=1}^{S} a_{sr} x_r(\mathbf{p}) - x_s(\mathbf{p}) \right), \quad (62)$$

where the last term in parenthesis is Home's imports in sector s, or

$$m_s(\mathbf{p}) \equiv c_s(p_s) + \sum\nolimits_{r=1}^{S} a_{sr} x_r(\mathbf{p}) - x_s(\mathbf{p}).$$

Welfare in the Foreign country can be expressed in an analogous manner.

We next consider the effects of Home levying a vector of trade taxes $\mathbf{t} = (t_{1,\ldots}, t_S)$ that generate a wedge between domestic and world prices. Notice that tariff levels (and associated price wedges $p_s - p^*_s$) are the same regardless of whether the good is being imported as a final good or as an intermediate input. Maximizing $W(\mathbf{p}, \mathbf{p}^*)$ with respect to t_s yields

$$\sum_{r=1}^{S} \left(\sum_{t=1}^{S} \frac{\partial \Pi_t(\mathbf{p})}{\partial p_r} + \frac{\partial S_r(p_r)}{\partial p_r} \right) \times \frac{\partial p_r}{\partial t_s} + \sum_{r=1}^{S} \left(\frac{\partial p_r}{\partial t_s} - \frac{\partial p^*_r}{\partial t_s} \right) \times m_r(\mathbf{p}) \\ + \sum_{r=1}^{S} (p_r - p^*_r) \sum_{t=1}^{S} \frac{\partial m_r(\mathbf{p})}{\partial p_t} \times \frac{\partial p_t}{\partial t_s} = 0.$$

Noting that $\partial S_r(p_r) / \partial p_r = -c_r(p_r)$, and that

$$\sum_{t=1}^{S} \frac{\partial \Pi_t(\mathbf{p})}{\partial p_r} = x_r(\mathbf{p}) - \sum\nolimits_{t=1}^{S} a_{rt} x_t(\mathbf{p}),$$

we can simplify the above expression to

$$-\sum_{r=1}^{S} \frac{\partial p^*_r}{\partial t_s} m_r(\mathbf{p}) + \sum_{r=1}^{S} (p_r - p^*_r) \sum_{t=1}^{S} \frac{\partial m_r(\mathbf{p})}{\partial p_t} \times \frac{\partial p_t}{\partial t_s} = 0.$$

Next noting that goods market clearing imposes $m_s(\mathbf{p}) = -m^*_r(\mathbf{p}^*)$, we can further simplify this to:

$$\sum_{r=1}^{S} \frac{\partial p^*_r}{\partial t_s} m^*_r(\mathbf{p}^*) = \sum_{r=1}^{S} (p_r - p^*_r) \sum_{t=1}^{S} \frac{\partial m^*_r(\mathbf{p}^*)}{\partial p^*_t} \frac{\partial p^*_t}{\partial t_s}. \quad (63)$$

If one were to ignore cross-price effects on import volumes, this formula would reduce to the familiar equation

$$t_s - 1 = \frac{1}{\varepsilon^*_{s,X}} \quad (64)$$

where

$$\varepsilon^*_{s,X} = \frac{\partial m^*_s(\mathbf{p}^*)}{\partial p^*_s} \frac{p^*_s}{m^*_s(\mathbf{p}^*)} \quad (65)$$

is the sectoral (partial-equilibrium) foreign export supply elasticity. Under this inverse elasticity formula in equation (64), whether optimal tariffs are higher or lower in different sectors would depend solely on this elasticity, regardless of other characteristics of these sectors (i.e., whether they are relatively upstream or downstream in GVCs).

Nevertheless, with vertical linkages across countries, there will naturally be cross-price effects on the supply side, which will generate cross-price effects on net import volumes, even when one shuts down cross-price effects on the demand side, as we have done with our assumption of separable quasi-linear preferences. We can of course still re-define (63) as

$$(d\mathbf{p})\left(m^{*}\left(\mathbf{p}^{*}\right)\right)^{T}=\left(\mathbf{p}-\mathbf{p}^{*}\right)\left(d\mathbf{m}^{*}\right)^{T},$$

which then equates tariffs with a 'general-equilibrium' export supply elasticity. In practice, however, this is not an elasticity that is straightforward to estimate, so this general formula provides little guidance in assessing whether the observed structure of protection is in line with the one maximizing social welfare. To make progress on this issue, we next turn to a special case of this more general model.

6.3 Tariffs on Final Goods and on Inputs in Competitive Economies

Let us consider the same environment as before, but assume now that there are only two sectors, other than the outside good sector 0. Sector F produces a good that is only consumed as a final good, while sector I produces a good that is only used as an input in production.[83] This simplified setting allows one to gain some insights on how intermediate input trade affects optimal final-good tariffs, and also what determines the optimal structure of protection in intermediate input sectors.

In terms of the model above, this is simply a special case with $S=\{F,I\}$, $S_I\left(p_I\right)=0$, and rent functions

$$\begin{aligned}\Pi_F\left(p_F,p_I\right) &= \left(p_F-ap_I\right)x_F-\ell_F\\ \Pi_I\left(p_I\right) &= p_Ix_I-\ell_I.\end{aligned}$$

Applying the formula in (63), and simplifying the system under the assumption that $\partial p_r^{*}/\partial t_s\neq 0$ for $s,r\in\{F,I\}$, we obtain

$$1 = \left(t_F-1\right)\varepsilon_{F,X}^{*}+\left(t_I-1\right)\frac{1}{m_F^{*}\left(\mathbf{p}^{*}\right)}\frac{\partial m_I^{*}\left(\mathbf{p}^{*}\right)}{\partial p_F^{*}} \qquad (66)$$

$$1 = \left(t_F-1\right)\frac{1}{m_I^{*}\left(\mathbf{p}^{*}\right)}\frac{\partial m_F^{*}\left(\mathbf{p}^{*}\right)}{\partial p_I^{*}}+\left(t_I-1\right)\varepsilon_{I,X}^{*}, \qquad (67)$$

where $\varepsilon_{F,X}^{*}$ and $\varepsilon_{I,X}^{*}$ are the sectoral foreign export supply elasticities.

As is clear from these expressions, the presence of vertical linkages introduces a deviation or wedge relative to the standard formula linking sectoral optimal tariffs to sectoral inverse foreign

---

[83] Other work studying optimal tariffs on final goods and inputs in neoclassical environments includes the work of Ruffin (1969), Casas (1973), and Das (1983).

export supply elasticities. What is the nature and sign of this wedge? To shed light on this, we first note that $m_F^*(\mathbf{p}^*) = c_F(\mathbf{p}^*) - x_F(\mathbf{p}^*)$ and $m_I^*(\mathbf{p}^*) = ax_F(\mathbf{p}^*) - x_I(\mathbf{p}^*)$ imply

$$\begin{aligned} \frac{\partial m_I^*(\mathbf{p}^*)}{\partial p_F^*} &= a\frac{\partial x_F(\mathbf{p}^*)}{\partial p_F^*} \geq 0 \\ \frac{\partial m_F^*(\mathbf{p}^*)}{\partial p_I^*} &= -\frac{\partial x_F^*(\mathbf{p}^*)}{\partial p_I^*} \geq 0. \end{aligned}$$

Intuitively, other things equal, an increase in the world price in sector F increases input demand in the RoW, thus increasing the RoW's net imports of inputs. Similarly, an increase in the world input price decreases the return to final-good production, decreasing RoW final-good output, and thus increasing the net imports of final goods.

Having signed these cross-effects, we can now return to equations (66) and (67) and note that the manner in which vertical linkages affect optimal tariffs depends on subtle aspects of the environment. To see this, note that if the Home country is importing goods F and I, then $m_F^*(\mathbf{p}^*) < 0$ and $m_I^*(\mathbf{p}^*) < 0$, and we necessarily have $t_F - 1 > 1/\varepsilon_{F,X}^*$ and $t_I - 1 > 1/\varepsilon_{I,X}^*$. The intuition is as follows: a final-good tariff provides the standard terms-of-trade gain, but in addition, by lowering the world price of the final good, it reduces the demand for intermediate inputs. This in turn reduces the world price of intermediate inputs, which affords an additional terms-of-trade gain, given that Home also imports inputs. Similarly, levying an import tariff on inputs not only affords the standard terms-of-trade gain, but the reduction in the world price of inputs, leads to an increase in the production of final-goods abroad, which reduces the world price of final goods as well, leading to an additional terms-of-trade gain when the Home country also imports final goods.

In sum, when Home imports both goods, vertical linkages lead to *higher* final-good tariffs and *higher* input tariffs than if these sectors were setting optimal tariffs based on the standard inverse-elasticity formula. Whether the wedge is higher or lower for final goods or for intermediates is less clear, however. With our assumption that inputs are used in fixed proportions, it turns out that $\partial m_I^*(\mathbf{p}^*)/\partial p_F^* = \partial m_F^*(\mathbf{p}^*)/\partial p_I^*$, and the model generates tariff escalation whenever (i) inverse export supply elasticities are (weakly) higher for final goods than for inputs, and (ii) net imports of intermediate inputs are higher than net imports of final goods.

The above results rely, however, on Home importing both final-goods and inputs in the industry under study. If, for instance, Home is a net exporter of inputs, we have $m_I^*(\mathbf{p}^*) > 0$, and thus, as long as $t_F > 1$, equation (67) implies that we must have $(t_I - 1)\varepsilon_{I,X}^* < 1$, and because $\varepsilon_{I,X}^* < 0$, this implies $0 > t_I - 1 > 1/\varepsilon_{I,X}^*$. Thus, Home sets a *lower* export tax on inputs than it would do under the standard formula. Now, from equation (66), $t_I < 1$ and $m_F^*(\mathbf{p}^*) < 0$ imply $(t_F - 1) < 1/\varepsilon_{F,X}^*$, and thus Home also sets a *lower* import tariff on final goods whenever Home is a net exporter of inputs. The intuition is as follows: by levying an import tariff on final goods, Home reduces the world price of final goods, thereby reducing the demand for inputs in the RoW. This puts downward pressure on the world price of inputs, which constitutes a terms-of-trade loss for Home.

Following analogous steps, one can show that Home will also set a lower export tax on final goods and a lower import tariff on inputs whenever Home exports final goods and imports inputs. And, finally, when Home is a net exporter of both goods, it will choose to set higher export taxes on both goods than it would under the standard inverse elasticity formula.

The bottom line of all this discussion is that even in a simple world with just two goods and an outside sector, how vertical linkages affect the level of optimal tariffs very much depends on the pattern of trade.

6.4 Political Economy, Lobbying Competition and Tariff Escalation

We next revert back to the roundabout model with general input-output links and consider the role of political economy forces in shaping the structure of protection in a world of GVCs. As microfounded in the work of Grossman and Helpman (1994), we posit that policy makers choose tariffs to maximize a 'policy support function' of the type

$$W\left(\mathbf{p},\mathbf{p}^{*}\right)+\lambda\sum_{s=1}^{S}\Pi_{s}\left(\mathbf{p}\right),$$

where $W\left(\mathbf{p},\mathbf{p}^{*}\right)$ is social welfare in equation (62), and $\lambda>0$ is an additional weight put on producer surplus. This latter term reflects the notion that concentration in the ownership of the sector-specific types of capital used in production allows producers to solve the collective action problem inherent in lobbying for protection, while consumers remain unorganized and ineffective in fighting producers' demands for protection. To isolate the role of political forces, we follow Grossman and Helpman (1994), Cadot et al. (2004), and Gawande et al. (2012), and assume that Home is a small-open economy. Under this assumption, and regardless of the structure on input-output links, a social planner would choose free trade in all sectors, as is clear from equation (63), where $p_s=p_s^{*}$ for all s whenever $\partial p_r^{*}/\partial t_s=0$ for all r and s.

In the presence of these political-economy forces, the system of first-order conditions characterizing optimal tariffs becomes

$$\sum_{r=1}^{S}\left(p_r-p_r^{*}\right)\frac{\partial m_r\left(\mathbf{p}\right)}{\partial p_s}\times\frac{\partial p_s}{\partial t_s}+\lambda\left(\sum_{t=1}^{S}\frac{\partial\Pi_t}{\partial p_s}\right)\frac{\partial p_s}{\partial t_s}=0.$$

Invoking the rent function in (61), we can further reduce this to

$$\sum_{r=1}^{S}\left(p_r-p_r^{*}\right)\frac{\partial m_r\left(\mathbf{p}\right)}{\partial p_s}+\lambda\left(x_s-\sum_{t=1}^{S}a_{sr}x_r\right)=0. \tag{68}$$

The interpretation of the two terms in equation (68) is as follows. The first term is analogous to the last term in the general expression (63), but it is somewhat simpler because one need not worry about tariff revenue effects working through changes in prices in other sectors. The second term is more novel and reflects the competition between sector-specific interests. As in Grossman and Helpman (1994), other things equal, sectors with higher output levels (relative to import volumes) will achieve higher protection, but note that this effect is attenuated by the usage of these sector's output in other sectors. Intuitively, the more a sector's output is used as an input in other sectors, the higher will be the cost of protection in that sector for other sectors in the economy. Because these other sectors have a voice in the political process (or, more broadly, in the process of tariff formation), this counterlobbying will tend to reduce the level of protection. As indicated by Cadot et al. (2004), this implies that, other things equal, relatively downstream sectors that sell predominantly to consumers should achieve higher levels of protection than relatively upstream sectors that sell predominantly to other sectors. The authors view this prediction as providing a rationale for the phenomenon of tariff escalation, which as described in Section 6.1, has been widely documented in the literature.

This result can be precisely formalized as follows. Suppose that the policy maker envisions levying a tariff in only one sector s, so $p_r = p_r^*$ for all $r \neq s$. Then equation (68) simplifies to

$$\frac{t_s - 1}{t_s} = \frac{\lambda}{\varepsilon_{s,M}} \frac{x_s - \sum_{t=1}^{S} a_{sr} x_r}{m_s},$$

and thus tariffs are larger the larger is λ, the lower the import demand elasticity $\varepsilon_{s,M} = -\left(\partial m_s / \partial p_s\right)\left(p_s / m_s\right)$, and the larger the ratio of final-good sales relative to total imports in the sector (rather than the standard import penetration ratio, as in the literature without input-output links).

When the policy maker chooses positive protection in various sectors, matters become more complicated due to the second-best effect of tariffs working through tariff revenue. More specifically, when raising a tariff t_s, output x_r in sectors using x_s as an input will be depressed, which will tend to increase imports in those sectors and generate higher tariff revenue, a force that works against the model producing tariff escalation.[84] Indeed, this is precisely the mechanism that leads Gawande and Bandyopadhyay (2000) and McCalman (2004) to derive a positive effect of upstream tariffs on downstream tariffs in models in which lobbying is *only* carried out by downstream producers. These same authors, as well as Erbahar and Zi (2017) more recently, provide empirical evidence supporting the relevance of this *cascading* trade protection.

Despite these conflicting effects, Cadot et al. (2004) perform numerical simulations of a variant of this model, and they show that the cross-industry tariff revenue term in (68) is quantitatively small, and thus the model delivers implications consistent with data. Further empirical evidence consistent with a model of lobby competition with upstream and downstream sectors is provided in Gawande et al. (2012). Using tariff, trade and input-output data from 42 countries at different levels of development, Gawande et al. (2012) show that country- and sector-specific tariffs are decreasing in the extent to which that country-sector's output is used as an input in other sectors. Furthermore, taking into account counterlobbying forces, leads to estimates of λ that are higher than when estimating models without such counterlobbying, such as Grossman and Helpman (1994)'s 'protection for sale' model (see Goldberg and Maggi, 1999). Intuitively, the standard model can only justify low tariffs via a low weight placed on lobbying contributions (which maps to λ above), and

---

[84] One way to shut down these effects – implicitly invoked by Gawande et al. (2012)– is to assume that technology combines sector-specific capital and labor in fixed proportions. In such case, sectoral output is fixed and independent of input or output prices in any sector.

thus a higher weight on social welfare. Meanwhile, the model with input-output links is consistent with a higher value of λ leading to lower tariffs on account of counterlobbying forces, and thus the implied benevolence of the policy maker is diminished.

6.5 Value-Added Approach

We next consider an environment inspired by the work of Blanchard et al. (2016). Although this is not essential for some of the results below, we stick to the competitive model developed in Section 6.3. The main novelty is that we now consider an environment in which production of final goods in each country combines labor, sector-specific capital, and an input that is *only* produced in the other country. From the point of view of the Home country, Blanchard et al. (2016) refer to the imported input as 'Foreign Value Added" (or FVA) used in Home production, and to the exported input as 'Domestic Value Added' (or DVA) in Foreign production.

Under these assumptions, the sector-specific capital at Home and in Foreign now earn

$$\begin{aligned}\Pi_F\left(p_F, p_{IFor}\right) &= p_F x_F - p_{IFor} x_{IFor} - \ell_F \\ \Pi_F^*\left(p_F^*, p_{IDom}^*\right) &= p_F^* x_F^* - p_{IDom}^* x_{IDom}^* - \ell_F^*,\end{aligned}$$

where x_{IFor} is the Foreign input used in Home production, and x_{IDom}^* is the Home input used in Foreign, and the corresponding prices for these inputs are p_{IFor} and p_{IDom}^*, respectively. For simplicity, we follow Blanchard et al. (2016) in assuming that these inputs are in fixed supply and remain untaxed, so the income obtained from selling these inputs is a pure rent given by

$$\begin{aligned}\Pi_I\left(p_{IDom}^*\right) &= p_{IDom}^* x_{IDom}^* \\ \Pi_I^*\left(p_{IFor}\right) &= p_{IFor} x_{IFor}.\end{aligned}$$

Welfare at Home is now given by

$$W\left(\mathbf{p}, \mathbf{p}^*\right) = 1 + \Pi_F\left(p_F, p_{IFor}\right) + \Pi_I\left(p_{IDom}^*\right) + S_F\left(p_F\right) + \left(p_F - p_F^*\right)\left(c_F\left(p_F\right) - x_F\left(p_F\right)\right),$$

where $\mathbf{p} = \left(p_F, p_{IFor}\right)$ and $\mathbf{p}^* = \left(p_F^*, p_{IDom}^*\right)$. Note that because inputs are in fixed supply, changes in the price of the Foreign input have no effect on Home's final-good production, and thus we can write $x_F\left(p_F\right)$.

We next consider the effects of levying an import tariff t_F on the final good. Differentiating $W\left(\mathbf{p}, \mathbf{p}^*\right)$ with respect to t_F and invoking $m_F\left(\mathbf{p}\right) = -m_F^*\left(\mathbf{p}\right)$ to simplify, delivers

$$t_F - 1 = \frac{1}{\varepsilon_{F,X}^*}\left(1 - \frac{p_{IFor} x_{IFor}}{p_F^* m_F} \xi_{For} - \frac{p_{IDom}^* x_{IDom}^*}{p_F^* m_F} \times \xi_{Dom}^*\right), \tag{69}$$

where $\varepsilon_{F,X}^*$ is the standard export supply elasticity defined in equation (65), and ξ_{For} and ξ_{Dom}^* are positive terms.[85]

Equation (69) makes it clear that there are two terms that generate a wedge between the optimal final-good tariffs in a standard model without input trade, and in this model featuring foreign value added in domestic production, and domestic value added in foreign production. These terms capture the intuitive nature that levying tariffs is more costly when (i) part of the rents obtained from that protection accrue to foreigners in the form of higher input prices p_{IFor}, and when (ii) the fall in the Foreign final-good price caused by the tariff reduces the rents that domestic value added obtains abroad (by reducing p^*_{IDom}). In sum, optimal final-good tariffs are predicted to be lower, the higher the foreign value added in domestic production (the term $p_{IFor}x_{IFor}$ in (69)), and the higher is the domestic value added in foreign production (the term $p^*_{IDom}x^*_{IDom}$ in (69)). Blanchard et al. (2016) explore various extensions of this framework that incorporate multiple sectors, political economy biases, preferences featuring income effects, and elastic supply of inputs. Their two key results continue to hold in those environments, though the expressions for optimal tariffs become significantly more complicated.

Blanchard et al. (2016) also explore the empirical validity of their theoretical predictions. To do so, they combine data on world input-output links from the WIOD, tariff data from World Bank's WITS website, as well as data on temporary trade barriers from Chad Bown's Temporary Trade Barriers Database (Bown, 2014). They find evidence in support of these two key predictions both when looking at how countries discriminate across trading partners by lowering protection through bilateral tariff preferences, and also when countries discriminate by raising protection through the adoption of temporary trade barriers, particularly against China.

It is also worth noting that the work of Blanchard et al. (2016) very much relates to prior insights from the work of Blanchard (2007), who showed that optimal tariffs tend to be lower in environments with foreign direct investment and international ownership, in which a country's factors deployed abroad suffer negative consequences from changes in terms of trade induced by tariffs. Another related work is Blanchard (2010), who studies the role of multilateral trade agreements in a world in which foreign investment leads to muted incentives for terms-of-trade manipulation.

6.6 Product Differentiation and General Equilibrium

So far, we have restricted the analysis to partial-equilibrium environments in which wages are pinned down by an outside sector and in which goods are homogeneous. This creates a bit of disconnect with modern macro models of GVCs, which as we have seen in Section 4, tend to be general-equilibrium in nature and generate bilateral gross exports and imports within sectors.

We next build on the recent work of Beshkar and Lashkaripour (2020), who consider optimal tariffs in a general equilibrium environment with roundabout production in which goods are differentiated by their country of production. Although their underlying economic model is significantly more general (see below), it is useful to focus attention on a discussion of their main results for the case in

---

[85] More specifically, $\xi_{For} \equiv \frac{(\partial p_{IFor}/\partial t_F)(t_F/p_{IFor})}{-(\partial p^*_F/\partial t_F)(t_F/p^*_F)}$ and $\xi^*_{Dom} \equiv \frac{-(\partial p^*_{IDom}/\partial t_F)(t_F/p^*_{IDom})}{-(\partial p^*_F/\partial t_F)(t_F/p^*_F)}$.

which their model reduces to the Caliendo and Parro (2015) framework. This amounts to assuming that (i) labor is the only factor of production, (ii) preferences are Cobb-Douglas across sectors, and CES across differentiated varieties within sectors, and (iii) technology is Cobb-Douglas in labor and the bundle of inputs in various sectors, with the latter being a CES aggregator of the differentiated inputs.

Building on the tools developed by Costinot et al. (2015), Beshkar and Lashkaripour (2020) solve for optimal trade taxes at the sectoral level, with the same tax levels applying regardless of the end use (i.e., final good or input) of the good being traded. Because, under product differentiation, countries will both export and imports varieties within sectors, Beshkar and Lashkaripour (2020) begin by considering a setting in which a government sets both optimal import tariffs t_s^M *and* optimal export taxes t_s^X against the Rest of the World. These optimal trade taxes are given by:

$$1 + t_s^M = 1 + \bar{t} \tag{70}$$

$$1 + t_s^X = \left(\frac{1 + \theta^s \pi_{ff}^s}{\theta^s \pi_{ff}^s - \sum_{r=1}^{s} \gamma_f^{sr} \xi_{fh}^r} \right) (1 + \bar{t})^{-1} \tag{71}$$

where $\bar{t}$ is an arbitrary constant, θ^s is the trade elasticity in sector s, π_{ff}^s is the share of spending by the Rest of the World on their own goods in sector s (defined in equation (16)), γ_f^{rs} is the (constant) share of production costs spent on intermediate inputs from sector s by industry-r producers in the Rest of the World, and ξ_{fh}^r is the share of output from the Rest of the World that is sold at Home (or $X_h^s \pi_{fh}^s / \left(X_h^s \pi_{fh}^s + X_f^s \pi_{ff}^s \right)$ using the notation developed in Section 4.1).

Several observations are in order regarding the optimal import and export taxes in equations (70) and (71). First, Lerner symmetry applies in this framework, and thus the level of optimal import and export taxes is only determined up to a common multiplicative shifter $(1 + \bar{t})$. Second, and in line with the results in Costinot et al. (2015), the presence of input-output linkages does not undo the fact that in Ricardian economies, optimal input tariffs are uniform across sectors. Intuitively, the main goal of import tariffs in this environment is to improve the countries terms of trade, and because all sector's technology is linear in labor, this implies that the incentive to mark-down imports is common in all sectors.[86] On the other hand, optimal export taxes are heterogeneous across sectors. The reason for this is that the main goal of export taxes is to exploit Home's market power in the differentiated varieties it produces. This leads to a markup (e.g., an export tax) that is negatively related to the trade elasticity θ^s and to the own trade share π_{ff}^s in the Rest of the World, but that is positively related to the extent to which a sector's output is sold to final consumers (low γ_f^{sr}), rather than being exported as an input and then reimported embedded in foreign goods (as captured by the term ξ_{fh}^r). The negative impact of $\theta^s \pi_{ff}^s$ on the optimal export tariff is well-understood, and is associated with the well-known formula developed by Gros (1987). The other effects are unique to the presence of intersectoral linkages in the model. In a nutshell, the use of export taxes will be curtailed in situations in which the higher export prices will be

---

[86] Beshkar and Lashkaripour (2020) consider an extension of their framework that incorporates sector-specific capital, and variation across sectors in the output elasticity of the specific capital generates variation in import tariffs.

passed-through back to Home consumers, when the goods being exported are largely inputs that are re-imported. As a result, Beshkar and Lashkaripour (2020) conclude that optimal export taxes are lower in upstream sectors than in downstream sectors, and they are also lower in a world of GVCs than in a world without vertical linkages across sectors.

Beshkar and Lashkaripour (2020) also consider the relevant case in which export taxes are ruled out.[87] In that case, import tariffs cease to be uniform across sectors, as they serve a second-best role in exploiting Home's market power in their export sectors. Consistently with the logic above, however, the incentive to levy import tariffs is higher in relatively upstream sectors, because the cost of the tariff is partly passed on to Foreign consumers in the form of higher input prices. As a result, their framework delivers a form of tariff de-escalation that is inconsistent with the observed negative correlation between import tariffs and upstreamness observed in the real world. This may be interpreted as indicating that the type of political economy effects emphasized by Cadot et al. (2004) and Gawande et al. (2012) are key for rationalizing tariff escalation practices. An alternative explanation, however, is that the pattern of optimal tariffs may look quite different under different market structures. We next turn to further explore the latter possibility.

6.7 Imperfect Competition

It is well understood that the study of trade policy outside the paradigm of perfectly competitive models is complicated by the fact that the nature and even the sign of optimal trade taxes is sensitive to details of how imperfect competition is modeled (see Eaton and Grossman, 1986). At the same time, the trade literature has largely converged to a particular approach to modeling imperfect competition – along the lines of Krugman (1980) – so it seems natural to explore optimal trade policy in versions of that environment that incorporate intermediate input trade. As in Krugman (1980), the focus is on an environment with CES preferences within sectors, increasing returns to scale technologies featuring constant marginal costs, and monopolistically competitive environments.

In carrying out this analysis, one could in principle follow a partial equilibrium approach or a general equilibrium approach, very much in line with the dichotomy outlined above in perfectly competitive environments. The partial equilibrium approach to the study of trade policy in monopolistically competitive environments originates in the work of Venables (1987), where the focus is solely on trade in final goods. The framework in Venables (1987) eliminates terms of trade effects working through wages by specifying an outside sector that pins down wages in all countries. The focus is instead on how trade taxes (and import tariffs, in particular) can enhance welfare by relocating (final-good) firm entry into one's own country, something that may prove beneficial for consumers – despite higher prices for imported goods – due to the presence of trade costs.[88] The general equilibrium approach to optimal trade policy with monopolistic competition – best exemplified by the work of Gros (1987) – instead re-focuses attention on terms-of-trade effects, and

---

[87] Export taxes are rarely used in practice, and some countries, such as the United States, explictly ban them (see U.S. Constitution, Article 1, Section 9, Clause 5).

[88] Ossa (2011) further developed Venables (1987)'s analysis and built a theory of trade agreements based on it.

emphasizes that small-open economies may have market power whenever the goods they import and export are differentiated. In terms of relocation effects, these are entirely eliminated in the one-sector model in Gros (1987).

A nascent literature is currently exploring variants of these models that feature intermediate input trade, and highlights the novel forces that arise in that case. Two recent examples are Caliendo et al. (2021) and Antràs et al. (2021), who both consider multi-sector Ricardian economies featuring general equilibrium terms-of-trade forces, but also Venables-style production relocation effects.

The simplest version of Caliendo et al. (2021)'s framework considers an environment analogous to that in Gros (1987) but with production being roundabout in nature. With imperfect competition, producers charge a markup over their marginal cost when selling intermediate inputs. Because, when selling to consumers, firms also charge a markup over marginal cost, the model features a double-marginalization inefficiency, that is absent in models without input trade. This inefficiency can be undone with targeted domestic subsidies, but in their absence, Caliendo et al. (2021) show that there is a second-best rationale for setting import tariffs that are lower than without roundabout production. The authors interpret this result as suggesting that the rise of GVCs and intermediate input trade puts downward pressure on the incentives of countries to unilaterally set tariffs on their trading partners. Caliendo et al. (2021) further extend their framework to include a second, nontradable sector, as well as firm heterogeneity à la Melitz (2003) and Chaney (2008). They demonstrate that despite the novel features that arise in that extension (such as the presence of relocation effects), their qualitative results still hold (and are even reinforced) in that expanded environment. In a quantitative exercise involving 186 countries, they also show that the magnitude of these mechanisms is sizeable, leading to a median optimal tariff of 10%, as compared to 27% in the absence of roundabout production.

Another recent contribution to this burgeoning literature is the work of Antràs et al. (2021). Rather than introducing intermediate input trade via roundabout production, these authors instead consider a two-sector version of the Krugman (1980) framework, with a novel intermediate-input sector which provides a bundle of differentiated input varieties to a Krugman-like final-good sector producing differentiated consumer varieties combining labor and a bundle of intermediate input varieties. The other assumptions of the model (a unique primitive factor of production, CES aggregators, increasing-returns-to-scale technologies, and monopolistic competition in both sectors) are all identical to those in Krugman (1980). This framework features the same type of double marginalization inefficiency highlighted by Caliendo et al. (2021), but the focus in Antràs et al. (2021) is instead to compare the differential incentives to levy import tariffs on final goods versus intermediate inputs. The equations characterizing these optimal tariffs are involved, but Antràs et al. (2021) develop tools to characterize the various mechanisms (terms of trade effects, production relocation effects, etc.) through which tariffs on final goods and inputs affect the general equilibrium of the model and welfare.[89] Their main result is that the optimal tariff is positive for both final

---

[89] In particular, they evaluate the first-order effect of increases in downstream and upstream tariffs starting from an equilibrium with zero tariffs. They also explore the size of optimal tariffs in both sectors by calibrating the model and performing quantitative exercises.

goods and for inputs, but the optimal tariff is higher for final goods than for inputs. Furthermore, the optimal tariff on inputs is quantitatively very small when domestic subsidies are available to undo the double marginalization inefficiency, while the optimal tariff on final goods remains high. Antràs et al. (2021) interpret their results as a potential rationale for observed tariff escalation practices.[90]

6.8 Trade Policy and Relational GVCs

We finally overview work that has studied the implications and design of trade policy in environments with the distinguishing features of relational GVCs, namely, search frictions, customized production, and incomplete contracting. Do these features introduce novel reasons for trade policy intervention? And do they create new problems of global policy cooperation motivating international agreements with novel features?

To study these questions, we begin by considering a simple framework inspired by the work of Antràs and Staiger (2012a). We consider a three-country world with two "small" countries, Home and Foreign, and a large Rest of the World (RoW). This large RoW pins down the untaxed world price of a single homogeneous final good, which is used as the numéraire. Production of the final good requires a customized input x, and technology is summarized by a production function $y(x)$, with $y(0) = 0$, $y'(x) > 0$, $y''(x) < 0$, $\lim_{x \to 0} y'(x) = +\infty$, and $\lim_{x \to \infty} y'(x) = 0$. Home only produces final goods, while Foreign only produces intermediate inputs. The marginal cost of input production in Foreign (measured in terms of the numéraire) is constant and, through choice of the units in which inputs are measured, it is normalized to 1. This implies that the efficient level of input production x^E absent any market imperfections, trade taxes or trade other barriers is implicitly characterized by $y'\left(x^E\right) = 1$.

It is assumed, however, that international contracts between suppliers and final-good producers are incomplete, and so the terms of trade between input suppliers and final good producers are determined by bargaining ex post after investments in input supply has already been sunk. For now, this is the only market friction we introduce. In particular, there is a unit measure of final-good producers at Home, and a unit measure of input producers in Foreign, and they are costlessly matched in pair. We shall consider environments with search frictions below.

We assume that each country can set trade taxes or subsidies on both the input and the final good. Because Foreign will never find it optimal to tax the final good, we can focus on the final-good tariff t_F^H at Home, defined in specific terms, and the input tariffs t_x^H and t_x^F set by Home and Foreign, also in specific terms. How do these instruments shape international exchanges? To explore this, note that domestic price of the final good at Home will be $1 + t_F^H$, while the input tariffs will increase the marginal cost of delivering inputs from 1 to $1 + t_x^H + t_x^F$. Recalling that the cost x of producing x units is sunk at the time the producer and supplier reach an agreement, the surplus these agents

---

[90] Antràs et al. (2021) also find that relocation effects are quantitatively as important for the determination of tariffs as standard terms-of-trade effects.

bargain over is given by

$$S\left(x, t_F^H, t_x^H, t_x^F\right) = t_F^H y\left(x\right) - \left(t_x^H + t_x^F\right) x. \tag{72}$$

We take the extreme view that inputs are completely customized to final-good producers, and that final-good producers have no recourse to a secondary market, so that the breakup of a bargaining pair would result in a zero outside option for both producer and supplier. It would be straightforward to relax this assumption along the lines of Ornelas and Turner (2008) or Antràs and Staiger (2012a). In the bargaining, we assume that final-good producers obtain a share β of the surplus $S\left(\cdot\right)$ in equation (72), while suppliers obtain the remaining share $1-\beta$. Before they reach the bargaining stage, suppliers will then set a level of investment $\tilde{x}$ that solves

$$\left(1-\beta\right)\left(1+\tau_1^H\right) y'\left(\tilde{x}\right) = 1 + \left(1-\beta\right)\left(t_x^H + t_x^F\right), \tag{73}$$

which implicitly defines $\tilde{x}(\tau_1^H, t_x^H, t_x^F)$. It is clear from (73) that $\tilde{x}$ is increasing in τ_1^H and decreasing in the sum of τ_x^H and τ_x^F. Intuitively, incomplete contracting leads to rent-sharing between the producer and supplier, and the latter's incentives to invest tend to be higher whenever the surplus from investment is higher, that is when τ_1^H is higher and when τ_x^H and τ_x^F is lower. Antràs and Staiger (2012a) show that the positive dependence of $\tilde{x}$ on τ_1^H and the negative dependence of $\tilde{x}$ on τ_x^H and τ_x^F hold for a variety of extensions of their framework featuring search frictions (see below), partial specificity, a secondary market for inputs, the existence of domestic input suppliers, two-sided hold up problems, and vertical integration, among others.

The result that input taxes – that is Home import tariffs on inputs or Foreign export taxes on inputs – reduce suppliers' investments is one of the key results in Ornelas and Turner (2008). They leverage this result and argue that trade liberalization, by reducing the hold-up problem faced by suppliers leads to increases in trade flows that are above and beyond the standard increases in trade volumes predicted by models without contractual frictions.[91]

Although it is clear that final-good tariffs ameliorate the hold up problem while input tariffs aggravate it, it is not clear from equation (73) which trade policies will be socially efficient, and which ones will be unilaterally optimal from the point of view of the Home and Foreign governments. To answer these questions, one needs to take into account the effect of these policies on consumer surplus, producer surplus and tax revenue in each of the two countries. Antràs and Staiger (2012a) carry out such an analysis. They first show that show that an appropriate choice of input trade subsidies, combined with free trade in final goods, can fully resolve the international hold-up problem and allow countries to attain the first-best. This is actually pretty straightforward to infer from equation (73). Note that setting $\tau_1^H = 0$ and $t_x^H + t_x^F = -\beta/\left(1-\beta\right)$, results in $y'\left(\tilde{x}\right) = 1$, and thus $\tilde{x} = x^E$. In sum, an appropriately chosen combination of input subsidies (provided by Home, Foreign or both) is sufficient to resolve the hold-up problem, leaving no role for final-good tariffs to

---

[91] Ornelas and Turner (2012) instead develop a framework in which input tariffs may ammeliorate the hold-up problem of domestic suppliers. Intuitively, by raising the domestic price for a generic version of the goods they produce, an input can strengthen the bargaining power of suppliers and lead them to invest more.

affect social welfare (remember that both Home and Foreign are small-open economies).

Antràs and Staiger (2012a) next show that the Nash equilibrium policy choices of governments do not coincide with the internationally efficient policies, and they identify two dimensions of international inefficiency that arise under Nash policies. A first dimension is an inefficiently low input trade volume. Intuitively, input trade policy serves a dual role in this environment. On the one hand, as indicated above, subsidies to the exchange of intermediate inputs can help restore the volume of input trade toward its efficient level. On the other hand, input trade taxes can be used to redistribute surplus across countries, thereby shifting some of the cost of intervention on to trading partners.[92] A second dimension of inefficiency relates to the incentive of the Home government to also distort trade in the *final* good away from its free-trade level in order to reduce the domestic final good price and further shift bargaining surplus from foreign input suppliers to home final good producers.

Antràs and Staiger (2012a) then study the role of trade agreements in closing the gap between socially efficient and Nash equilibrium trade policy choices. Their key result is that the "shallow" focus on negotiations over market access advocated by the traditional terms-of-trade theory, is not effective in their setting. Instead, in the presence of bilaterally negotiated prices, effective trade agreements and the institutions that support them will have to evolve, from a market access focus toward a focus on deep integration, and from a reliance on simple and broadly-applied rules such as reciprocity and non-discrimination toward a collection of more-individualized agreements that can better reflect member-specific idiosyncratic needs. These lessons very much resonate with the empirical results of Orefice and Rocha (2014) and Laget et al. (2020), documenting a relationship between the increased prevalence of deep integration (as measured by the number of policy areas covered in preferential trade agreements) and the rise of GVCs.[93]

Despite the above emphasis on hold-up inefficiencies, Antràs and Staiger (2012b) clarified that the key aspect of the Antràs and Staiger (2012a) framework that delivers the need for "deeper" agreement is the fact that prices in international exchanges are not fully disciplined by market clearing conditions, and are instead bargained over. They further demonstrate that their main insights emerge in a much simpler framework without ex-ante investments but with search frictions in the matching between buyers and sellers. These matching frictions generate a lock-in effect that again leads to bargaining in international exchanges, with the resulting prices again not being fully disciplined by market clearing conditions.

The interplay between search frictions and trade policy takes a prominent role in the recent work by Grossman and Helpman (2020). Their framework shares some features with the work of Ornelas and Turner (2008), Antràs and Staiger (2012a) and Antràs and Staiger (2012b), but it is much richer in many dimensions, and it focuses on a different set of issues. The main goal of

---

[92] For instance, although an export tax may reduce the incentive of foreign suppliers to invest, these suppliers will be able to pass part of the cost of the tax on to Home final-good producers in their ex-post bargaining.

[93] Deep integration is often associated with provisions related to 'behind-the-border' trade barriers such as competition policy, labor market regulation, consumer protection, environmental laws, movement of capital, or intellectual property rights protection.

Grossman and Helpman (2020) is to study how input tariffs affect (i) the incentives of final-good producers to search for suppliers in various countries, and (ii) the bargaining between final-good producers and their various suppliers. In their framework, final-good producers assemble a bundle of intermediate inputs, and thus need to match with a continuum of suppliers. The productivity of these suppliers is in turn heterogeneous, but final-good producers do not learn that productivity until they are matched with a supplier. Because search costs are sunk in nature and the productivity of alternative matches is unknown, existing GVC relationships tend to be sticky in the model. As a result, relatively small unanticipated tariff changes do not lead to relocation of production, though in this case the tariffs can still cause contract renegotiations within existing matches due to the altered outside options of buyers in the presence of the tariffs. Large tariffs can overcome this stickiness and lead to the destruction of existing matches and the creation of new ones, either in the domestic economy or in other foreign countries.[94] The effects of input tariffs are ambiguous in their model and depend on various primitive parameters in rich ways, but numerical simulations of their model suggest that input tariffs are generally welfare reducing.

7 Conclusion and Future Directions

Over the past few decades, production processes have become increasingly more complex in the world economy. Any finished good now typically embodies value added from multiple countries of origin, with this value added often crossing multiple borders en route to its point of consumption, in production arrangements that have come to be referred to as "global value chains". In this article, we have surveyed recent developments and contributions by economists, particularly in the field of international trade, towards deepening our understanding of global value chains. At the same time, we have sought to identify various directions for continued research effort.

On the empirical side, we have reviewed efforts to measure GVCs as a "macro" phenomenon. Work on this front has advanced in recent years with improvements in World Input-Output Tables (WIOTs) and the development of value added accounting methodologies. This has been accompanied by a parallel large body of empirical work on firms in international trade; more specifically, this has uncovered useful facts at the "micro" level on selection into participation in GVCs, the formation of buyer-supplier links, and the relational nature of these ties.

On the theory side, we have attempted to organize the various modeling frameworks that address the phenomenon of global production. This includes "macro" models that incorporate rich input-output linkages across countries and industries, either in the form of a "roundabout" production setup or in more sequential production chains. Such models have provided valuable quantitative insights on the aggregate consequences of GVCs, relative to a world in which such input-output linkages are absent. On the other hand, a rich vein of "micro" models has shone the spotlight on the firm-level drivers of forward and backward participation in GVCs, as well as the relational aspects of these GVC links. Last but not least, we have surveyed a nascent body of work

---

[94] Grossman and Helpman (2020) also provide suggestive evidence of these relocation effects following the recent imposition of tariffs (many of them on inputs) by President Trump's administration.

seeking to understand the trade policy implications of a world with active trade in intermediate inputs within GVCs.

References

Acemoglu, D., P. Antràs, and E. Helpman (2007). Contracts and technology adoption. *American Economic Review 97*(3), 916–943.

Adao, R., P. Carrillo, A. Costinot, D. Donaldson, and D. Pomeranz (2020). International trade and earnings inequality: A new factor content approach. NBER Working Paper No. 28263.

Aichele, R., G. Felbermayr, and I. Heiland (2016). Going deep: The trade and welfare effects of TTIP revised. *Ifo-Working Paper 219*.

Aichele, R. and I. Heiland (2018). Where is the value added? trade liberalization and production networks. *Journal of International Economics 115*, 130–144.

Alexander, P. (2017). Vertical specialisation and gains from trade. Bank of Canada.

Alfaro, L., P. Antràs, D. Chor, and P. Conconi (2019). Internalizing Global Value Chains: A Firm-Level Analysis. *Journal of Political Economy 127*(2), 508–559.

Alfaro-Urena, A., I. Manelici, and J. P. Vasquez (2020). The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages. Mimeo.

Allen, T. (2014). Information frictions in trade. *Econometrica 82*(6), 2041–2083.

Altomonte, C., L. Bonacorsi, and I. Colantone (2018). Trade and Growth in the Age of Global Value Chains. Mimeo Bocconi University.

Amador, J. and F. Di Mauro (2015). The age of global value chains. *VOX CEPR's Policy Portal*.

Amiti, M., S. H. Kong, and D. Weinstein (2020). The Effect of the U.S.-China Trade War on U.S. Investment. NBER Working Paper No. 27114.

Amiti, M. and J. Konings (2007). Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia. *American Economic Review 97*(5), 1611–1638.

Amiti, M., S. J. Redding, and D. E. Weinstein (2019). The impact of the 2018 tariffs on prices and welfare. *Journal of Economic Perspectives 33*(4), 187–210.

Anderson, J. E. (1998). Effective protection redux. *Journal of international economics 44*(1), 21–44.

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

Antràs, P. (2005). Incomplete contracts and the product cycle. *American economic review 95*(4), 1054–1073.

Antràs, P. (2016). *Global Production: A Contracting Perspective*. Princeton, NJ: Princeton University Press.

Antràs, P. (2020). Conceptual aspects of global value chains. *World Bank Economic Review*, forthcoming.

Antràs, P. (2020). De-Globalisation? Global Value Chains in the Post-COVID-19 Age. NBER Working Paper No. 28115.

Antràs, P. and D. Chor (2013). Organizing the Global Value Chain. *Econometrica 81*(6), 2127–2204.

Antràs, P. and D. Chor (2019). On the Measurement of Upstreamness and Downstreamness in Global Value Chains. In L. Y. Ing and M. Yu (Eds.), *World Trade Evolution: Growth, Productivity and Employment*, Chapter 5, pp. 126–194. Routledge.

Antràs, P., D. Chor, T. Fally, and R. Hillberry (2012). Measuring the Upstreamness of Production and Trade Flows. *American Economic Review Papers and Proceedings 102*(3), 412–416.

Antràs, P. and A. Costinot (2011). Intermediated trade. *The Quarterly Journal of Economics 126*(3), 1319–1374.

Antràs, P. and A. de Gortari (2016). Global value chains, external economies to scale and the traveling salesman problem. unpublished notes.

Antràs, P. and A. de Gortari (2020). On the Geography of Global Value Chains. *Econometrica 84*(4), 1553–1598.

Antràs, P., E. Fadeev, T. Fort, and F. Tintelnot (2020). Global sourcing and multinational activity: A unified approach. Working Paper, Harvard University.

Antràs, P., E. Fadeev, A. Gutiérrez, T. Fort, and F. Tintelnot (2021). Import tariffs and global sourcing. Working Paper, Harvard University.

Antràs, P. and C. F. Foley (2015). Poultry in motion: A study of international trade finance practices. *Journal of Political Economy 123*(4), 853–901.

Antràs, P., T. Fort, and F. Tintelnot (2017). The Margins of Global Sourcing: Theory and Evidence from U.S. Firms. *American Economic Review 107*(9), 2514–2564.

Antràs, P., L. Garicano, and E. Rossi-Hansberg (2006). Offshoring in a knowledge economy. *The Quarterly Journal of Economics 121*(1), 31–77.

Antràs, P. and E. Helpman (2004). Global sourcing. *Journal of political Economy 112*(3), 552–580.

Antràs, P. and E. Helpman (2008). Contractual frictions and global sourcing. In E. Helpman, T. Verdier, and D. Marin (Eds.), *The Organization of Firms in a Global Economy*, Chapter 1, pp. 9–54. Cambridge, MA: Harvard University Press.

Antràs, P. and R. Staiger (2012a). Offshoring and the Role of Trade Agreements. *American Economic Review 102*(7), 3140–3183.

Antràs, P. and R. W. Staiger (2012b). Trade Agreements and the Nature of Price Determination. *American Economic Review Papers and Proceedings 102*(3), 470–476.

Antràs, P. and S. R. Yeaple (2014). Multinational firms and the structure of international trade. In *Handbook of International Economics*, Volume 4, pp. 55–130. Elsevier.

Araujo, L., G. Mion, and E. Ornelas (2016). Institutions and export dynamics. *Journal of International Economics 98*, 2–20.

Arkolakis, C. (2010). Market penetration costs and the new consumers margin in international trade. *Journal of Political Economy 118*(6), 1151–1199.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012). New trade models, same old gains. *American Economic Review 102*(1), 94–130.

Arkolakis, C. and F. Eckert (2017). Combinatorial Discrete Choice. *Working Paper*.

Arkolakis, C., N. Ramondo, A. Rodríguez-Clare, and S. Yeaple (2018). Innovation and Production in the Global Economy. *American Economic Review 108*(8), 2128–73.

Atalay, E., A. Hortaçsu, J. Roberts, and C. Syverson (2011). Network Structure of Production. *PNAS 108*(13), 5199–5202.

Autor, D., D. Dorn, and G. Hanson (2013). The China Syndrome: Local Labor Effects of Import Competition in the United States. *American Economic Review 103*(6), 2121 – 2168.

Bagwell, K., R. W. Staiger, and A. Yurukoglu (2018). Quantitative Analysis of Multi-Party Tariff Negotiations. NBER Working Paper No. 24273.

Baker, G., R. Gibbons, and K. J. Murphy (2002). Relational Contracts and the Theory of the Firm. *Quarterly Journal of Economics 117*(1), 39–84.

Balassa, B. (1965). Tariff protection in industrial countries: an evaluation. *Journal of Political Economy 73*(6), 573–594.

Balassa, B. (1971). *The Structure of Protection in Developing Countries.* John Hopkins Press.

Baldwin, R. (2016). *The Great Convergence.* Harvard University Press.

Baldwin, R. and A. Venables (2013). Spiders and Snakes: Offshoring and Agglomeration in the Global Economy. *Journal of International Economics 90*(2), 245–254.

Baqaee, D. and E. Farhi (2019). Networks, barriers, and trade. NBER Working Paper No. 26108.

Barattieri, A. and M. Cacciatore (2020). Self-harming trade policy? protectionism and production networks. NBER Working Paper No. 27630.

Barrot, J.-N. and J. Sauvagnat (2016). Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks. *Quarterly Journal of Economics 131*(3), 1543–1592.

Bas, M. (2012). Input-trade Liberalization and Firm Export Decisions: Evidence from Argentina. *Journal of International Economics 97*(2), 481–493.

Bas, M. and V. Strauss-Kahn (2014). Does Importing More Inputs Raise Exports? Firm-level Evidence from France. *Review of World Economics 150*, 241–275.

Bems, R. and R. C. Johnson (2017). Demand for Value Added and Value-Added Exchange Rates. *American Economic Journal: Macroeconomics 9*(4), 45–90.

Bems, R. and A. K. Kikkawa (2021). Measuring Trade in Value Added with Firm-level Data. *Journal of International Economics forthcoming*, forthcoming.

Benguria, F. (2021). The Matching and Sorting of Exporting and Importing Firms: Theory and Evidence. *Journal of International Economics forthcoming.*

Bernard, A., E. Bøler, and S. Dhingra (2019). Firm-to-Firm Connections in Colombian Imports. In L. Y. Ing and M. Yu (Eds.), *World Trade Evolution: Growth, Productivity and Employment*, Chapter 10, pp. 333–365. Routledge.

Bernard, A., E. Dhyne, G. Magerman, K. Manova, and A. Moxnes (2019). The origins of firm heterogeneity: A production network approach. NBER Working Paper No. 25441.

Bernard, A. B. and J. B. Jensen (1999). Exceptional Exporter Performance: Cause, Effect, or Both? *Journal of International Economics 47*(1), 1–25.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2007). Firms in International Trade. *Journal of Economic Perspectives 21*(3), 105–130.

Bernard, A. B., J. B. Jensen, S. R. Redding, and P. K. Schott (2018). Global firms. *Journal of Economic Literature 56*(2), 565–619.

Bernard, A. B., J. B. Jensen, and P. K. Schott (2009). Importers, Exporters and Multinationals: A Portrait of Firms in the U.S. that Trade Goods. In T. Dunne, J. B. Jensen, and M. J. Roberts (Eds.), *Producer Dynamics: New Evidence from Micro Data* , Chapter 14, pp. 513–552. University of Chicago Press.

Bernard, A. B. and A. Moxnes (2018). Networks and trade. *Annual Review of Economics 10*, 65–85.

Bernard, A. B., A. Moxnes, and Y. U. Saito (2019). Production networks, geography, and firm performance. *Journal of Political Economy 127*(2), 639–688.

Bernard, A. B., A. Moxnes, and K. H. Ulltveit-Moe (2018). Two-Sided Heterogeneity and Trade. *Review of Economics and Statistics 100*(3), 424–439.

Bernard, A. B., S. J. Redding, and P. K. Schott (2010). Multiple-Product Firms and Product Switching. *American Economic Review 100*(1), 70–97.

Bernard, A. B., S. J. Redding, and P. K. Schott (2011). Multiproduct firms and trade liberalization. *Quarterly Journal of Economics 126*(13), 1271–1318.

Besedeš, T. and T. J. Prusa (2006a). Ins, Outs, and the Duration of Trade. *Canadian Journal of Statistics 39*(1), 266–295.

Besedeš, T. and T. J. Prusa (2006b). Product Differentiation and the Duration of US Import Trade. *Canadian Journal of Statistics 70*(2), 339–358.

Besedeš, T. and T. J. Prusa (2011). The Role of Extensive and Intensive Margins and Export Growth. *Journal of International Economics 96*(2), 371–379.

Beshkar, M. and A. Lashkaripour (2020). The Cost of Dissolving the WTO: The Role of Global Value Chains. mimeo Indiana University.

Bilgin, N. M. (2020). Endogenous production networks and learning-by-networking. mimeo Bocconi University.

Blanchard, E. J. (2007). Foreign Direct Investment, Endogenous Tariffs, and Preferential Trade Agreements. *B.E. Journal of Economic Analysis and Policy - Advances 7*(1), Article 54.

Blanchard, E. J. (2010). Reevaluating the Role of Trade Agreements: Does Investment Globalization Make the WTO Obsolete? *Journal of International Economics 82(1)*, 63–72.

Blanchard, E. J., C. P. Bown, and R. C. Johnson (2016). Global supply chains and trade policy. NBER Working Paper 21883.

Blaum, J., C. Lelarge, and M. Peters (2018). The Gains from Input Trade with Heterogeneous Importers. *American Economic Journal: Macroeconomics 10*(4), 77–127.

Blaum, J., C. Lelarge, and M. Peters (2019). Firm Size, Quality Bias and Import Demand. *Journal of International Economics 120*, 59–83.

Blum, B. S., S. Claro, and I. Horstmann (2010). Facts and Figures on Intermediated Trade. *American Economic Review Papers & Proceedings 100*(2), 419–423.

Board, S. (2011). Relational contracts and the value of loyalty. *American Economic Review 101*(7), 3349–67.

Boehm, C. E., A. Flaaen, and N. Pandalai-Nayar (2019). Input Linkages and the Transmission of Shocks: Firm Level Evidence from the 2011 Tohoku Earthquake. *Review of Economics and Statistics 101*(1), 60–75.

Boehm, J. (2020). The Impact of Contract Enforcement Costs on Value Chains and Aggregate Productivity. *Review of Economics and Statistics forthcoming*, forthcoming.

Boehm, J., S. Dhingra, and J. Morrow (2018). Product Diversification in Indian Manufacturing. In W. Kohler and E. Yalcin (Eds.), *Developments in Global Sourcing*, Chapter 13, pp. 337–350. MIT Press.

Boehm, J. and E. Oberfield (2020). Misallocation in the Market for Inputs: Enforcement and the Organization of Production. *Quarterly Journal of Economics 135*(4), 2007–2058.

Bolatto, S., A. Naghavi, G. Ottaviano, and K. Z. Kejžar (2018). Intangible Assets and the Organization of Global Supply Chains. Mimeo.

Bøler, E. A., A. Moxnes, and K. H. Ulltveit-Moe (2015). R&D, International Sourcing, and the Joint Impact on Firm Performance. *American Economic Review 105*(12), 3704–3739.

Bonadio, B., Z. Huo, A. A. Levchenko, and N. Pandalai-Nayar (2020). Global supply chains in the pandemic. NBER Working Paper 27224.

Borin, A. and M. Mancini (2019). Measuring What Matters in Global Value Chains and Value-Added Trade. World Bank Policy Research Working Paper 8804.

Bown, C., P. Conconi, A. Erbahar, and L. Trimarchi (2020). Trade Protection Along Supply Chains. mimeo.

Bown, C. and M. Crowley (2016). Chapter 1 - the empirical landscape of trade policy. Volume 1 of *Handbook of Commercial Policy*, pp. 3 – 108. North-Holland.

Bown, C. and E. Zhang (2019, February). Measuring trump's 2018 trade protection: Five takeaways. *Peterson Institute Trade and Investment Policy Watch*.

Bown, C. P. (2014). Temporary Trade Barriers Database. http://econ.worldbank.org/ttbd/.

Brugués, F. (2020). Take the Goods and Run: Contracting Frictions and Market Power in Supply Chains. Brown University mimeo.

Burress, D. (1994). Homeomorphism between leontief and cobb-douglas input-output models. *Economics Letters 44*(1-2), 49–53.

Caceres, C., D. A. Cerdeiro, and R. Mano (2019). Trade wars and trade deals: Estimated effects using a multi-sector model. International Monetary Fund.

Cadot, O., J. De Melo, and M. Olarreaga (2004). Lobbying, counterlobbying, and the structure of tariff protection in poor and rich countries. *The World Bank Economic Review 18*(3), 345–366.

Cajal-Grossi, J., R. Macchiavello, and G. Noguera (2020). Buyers' Sourcing Strategies and Suppliers' Markups in Bangladeshi Garments. LSE mimeo.

Caliendo, L., M. Dvorkin, and F. Parro (2019). Trade and labor market dynamics: General equilibrium analysis of the China trade shock. *Econometrica 87*(3), 741–835.

Caliendo, L., R. C. Feenstra, J. Romalis, and A. M. Taylor (2021). A second-best argument for low optimal tariffs. NBER Working Paper 28380.

Caliendo, L. and F. Parro (2015). Estimates of the Trade and Welfare Effects of NAFTA. *Review of Economic Studies 82*(1), 1–44.

Caliendo, L., F. Parro, E. Rossi-Hansberg, and P.-D. Sarte (2018). The impact of regional and sectoral productivity changes on the US economy. *The Review of economic studies 85*(4), 2042–2096.

Caliendo, L., F. Parro, and A. Tsyvinski (2017). Distortions and the structure of the world economy. NBER Working Paper 23332.

Carballo, J., G. I. Ottaviano, and C. V. Martincus (2018). The buyer margins of firms' exports. *Journal of International Economics 112*, 33–49.

Carluccio, J. and M. Bas (2015). The impact of worker bargaining power on the organization of global firms. *Journal of International Economics 96*(1), 162–181.

Carrico, C., E. Corong, and D. van der Mensbrugghe (2020). The GTAP 10A Multi-Region Input Output (MRIO) Data Base. Mimeo.

Carvalho, V. M., M. Nirei, Y. U. Saito, and A. Tahbaz-Salehi (2020). Supply Chain Disruptions: Evidence from the Great East Japan Earthquake. *Quarterly Journal of Economics*, forthcoming.

Casas, F. R. (1973). Optimal effective protection in general equilibrium. *The American Economic Review 63*(4), 714–716.

Caselli, F., M. Koren, M. Lisicky, and S. Tenreyro (2020). Diversification through trade. *The Quarterly Journal of Economics 135*(1), 449–502.

Castellani, D., F. Serti, and C. Tomasi (2010). Firms in International Trade: Importers' and Exporters' Heterogeneity in Italian Manufacturing Industry. *World Economy 33*(3), 424–457.

Chaney, T. (2008). Distorted gravity: the intensive and extensive margins of international trade. *American Economic Review 98*(4), 1707–21.

Chaney, T. (2014). The network structure of international trade. *American Economic Review 104*(11), 3600–3634.

Chaney, T. (2018). The gravity equation in international trade: An explanation. *Journal of Political Economy 126*(1), 150–177.

Charbonneau, K. B. and A. Landry (2018). The trade war in numbers. Bank of Canada Staff Working Paper.

Charoenwong, B., M. Han, and J. Wu (2021). Not Coming Home: Trade and Economic Policy Uncertainty in American Supply Chain Networks. mimeo.

Chor, D. (2019). Modeling global value chains: approaches and insights from economics. In *Handbook on Global Value Chains*. Edward Elgar Publishing.

Chor, D. and L. Ma (2020). Contracting frictions in global sourcing: Implications for welfare. Technical Report, Dartmouth November.

Chor, D., K. Manova, and Z. Yu (2020). Growing Like China: Firm Performance and Global Production Line Position. NBER Working Paper No. 27795.

Clerides, S., 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.

Colantone, I. and R. Crinò (2014). New Imported Inputs, New Domestic Products. *Journal of International Economics 92*, 147–165.

Conconi, P., M. García-Santana, L. Puccio, and R. Venturini (2018). From final goods to inputs: the protectionist effect of rules of origin. *American Economic Review 108*(8), 2335–65.

Corden, W. M. (1966). The structure of a tariff system and the effective protective rate. *Journal of Political Economy 74*(3), 221–237.

Costinot, A., D. Donaldson, J. Vogel, and I. Werning (2015). Comparative advantage and optimal trade policy. *The Quarterly Journal of Economics 130*(2), 659–702.

Costinot, A. and A. Rodríguez-Clare (2014). Trade theory with numbers: Quantifying the consequences of globalization. In G. Gopinath, E. Helpman, and K. Rogoff (Eds.), *Handbook of International Economics*, Volume 4, Chapter 4. Amsterdam: Elsevier.

Costinot, A., J. Vogel, and S. Wang (2013). An Elementary Theory of Global Supply Chains. *Review of Economic Studies 80*(1), 109–144.

Das, S. P. (1983). Optimum tariffs on final and intermediate goods. *International Economic Review*, 493–508.

Dasgupta, K. and J. Mondria (2018). Inattentive importers. *Journal of International Economics 112*, 150–165.

Davis, D. R. and D. E. Weinstein (2001). An Account of Global Factor Trade. *American Economic Review 91*, 1423–1453.

de Gortari, A. (2019). Disentangling Global Value Chains. NBER Working Paper No. 25868.

de Gortari, A. (2020). Global Value Chains and Increasing Returns. Working Paper, Dartmouth College.

de Loecker, J. (2007). Do Exports Generate Higher Productivity? Evidence from Slovenia. *Journal of International Economics 73*(1), 69–98.

De Soyres, F., A. Mulabdic, S. Murray, N. Rocha, and M. Ruta (2018). *How much will the Belt and Road Initiative reduce trade costs?* The World Bank.

Dedrick, J., K. L. Kraemer, and G. Linden (2010). Who Profits from Innovation in Global Value Chains? A Study of the iPod and Notebook PCs. *Industrial and Corporate Change 19*(1), 81–116.

Defever, F., C. Fischer, and J. Suedekum (2016). Relational contracts and supplier turnover in the global economy. *Journal of International Economics 103*, 147–165.

Dekle, R., J. Eaton, and S. Kortum (2008). Global rebalancing with gravity: Measuring the burden of adjustment. *IMF Staff Papers 55*(3), 511–540.

del Prete, D. and A. Rungi (2017). Organizing the Global Value Chain: A Firm-Level Test. *Journal of International Economics 109*, 16–30.

Demir, B., A. C. Fieler, D. Y. Xu, and K. K. Yang (2021). O-Ring Production Networks. mimeo.

Dhingra, S., H. Huang, G. Ottaviano, J. Paulo Pessoa, T. Sampson, and J. Van Reenen (2017). The costs and benefits of leaving the eu: trade effects. *Economic Policy 32*(92), 651–705.

Dhyne, E., K. Kikkawa, M. Mogstad, and F. Tintelnot (2020). Trade and domestic production networks. *Review of Economic Studies*.

Di Giovanni, J., A. A. Levchenko, and J. Zhang (2014). The global welfare impact of China: Trade integration and technological change. *American Economic Journal: Macroeconomics 6*(3), 153–83.

Dietzenbacher, E., B. Los, R. Stehrer, M. P. Timmer, and G. J. de Vries (2013). The Construction of World Input-Output Tables in the WIOD Project. *Review of International Economics 25*(1), 71–98.

Dietzenbacher, E., I. R. Luna, and N. S. Bosma (2005). Using Average Propogration Lengths to Identify Production Chains in the Andalusian Economy. *Estudios de Economia Aplicada 23*(2), 405–422.

Dingel, J. I. and F. Tintelnot (2020). Spatial economics for granular settings. NBER Working Paper No. 27287.

Dixit, A. K. and G. M. Grossman (1982). Trade and protection with multistage production. *The Review of Economic Studies 49*(4), 583–594.

Dornbusch, R., S. Fischer, and P. A. Samuelson (1977). Comparative advantage, trade, and payments in a ricardian model with a continuum of goods. *The American Economic Review 67*(5), 823–839.

Eaton, J., M. Eslava, C. Krizan, M. Kugler, and J. Tybout (2014). A search and learning model of export dynamics. mimeo Pennsylvania State University.

Eaton, J. and G. M. Grossman (1986). Optimal trade and industrial policy under oligopoly. *The Quarterly Journal of Economics 101*(2), 383–406.

Eaton, J., D. Jinkins, J. Tybout, and D. Xu (2016). Two-sided search in international markets. mimeo Pennsylvania State University.

Eaton, J. and S. Kortum (2002). Technology, Geography, and Trade. *Econometrica 70*(5), 1741–1780.

Eaton, J., S. Kortum, and F. Kramarz (2018). Firm-to-firm trade: Imports, exports, and the labor market. mimeo Yale University.

Eppinger, P., G. Felbermayr, O. Krebs, and B. Kukharskyy (2020). Covid-19 shocking global value chains. CESifo Working Paper No. 8572.

Eppinger, P. and B. Kukharskyy (2020). Contracting institutions and firm integration around the world. *Working paper, University of Täubingen and City University of New York*.

Erbahar, A. and Y. Zi (2017). Cascading trade protection: Evidence from the us. *Journal of International Economics 108*, 274–299.

Ethier, W. (1977). The theory of effective protection in general equilibrium: Effective-rate analogues of nominal rates. *Canadian Journal of Economics*, 233–245.

Fajgelbaum, P. D., P. K. Goldberg, P. J. Kennedy, and A. K. Khandelwal (2020). The return to protectionism. *Quarterly Journal of Economics 135*(1), 1–55.

Fally, T. (2012). On the Fragmentation of Production in the U.S. mimeo, University of Berkeley.

Fally, T. and R. Hillberry (2018). A Coasian Model of International Production Chains. *Journal of International Economics 114*, 299–315.

Fally, T. and J. Sayre (2018). Commodity trade matters. NBER Working Paper 24965.

Fan, H., E. L.-C. Lai, and H. S. Qi (2019). Trade liberalization and firms' export performance in China: Theory and evidence. *Journal of Comparative Economics 47*(3), 640–668.

Fan, H., T. A. Luong, E. Lai, and L. Zhang (2021). Import liberalization and export product mix. *Canadian Journal of Economics*.

Farrokhi, F. (2020). Global sourcing in oil markets. *Journal of International Economics 125*, 103323.

Feenstra, R. (1998). Integration of Trade and Disintegration of Production in the Global Economy. *Journal of Economic Perspectives 12*(4), 31–50.

Feenstra, R. C. and G. H. Hanson (1996). Foreign Investment, Outsourcing and Relative Wages. In R. C. Feenstra, G. M. Grossman, and D. A. Irwin (Eds.), *Political Economy of Trade Policy: Essays in Honor of Jagdish Bhagwati*, Chapter , pp. 89–127. Cambridge: MIT Press.

Feng, L., Z. Li, and D. L. Swenson (2016). The Connection between Imported Intermediate Inputs and Exports: Evidence from Chinese Firms. *Journal of International Economics 101*, 86–101.

Fernandes, A., H. L. Kee, and D. Winkler (2020). Determinants of Global Value Chain Participation: Cross-Country Evidence. World Bank Policy Research Working Paper No. 9197.

Flaaen, A., A. Hortaçsu, and F. Tintelnot (2020). The Production Relocation and Price Effects of U.S. Trade Policy: The Case of Washing Machiness. *American Economic Review 110*(7), 2103–2127.

Flaaen, A. and J. Pierce (2019). Disentangling the Effects of the 2018-2019 Tariffs on a Globally Connected US Manufacturing Sector. Finance and Economics Discussion Series Working Paper No. 086.

Furusawa, T., T. Inui, K. Ito, and H. Tang (2017). Global sourcing and domestic production networks. CESifo Working Paper Series No. 6658.

Furusawa, T. and Y. Sugita (2020). The welfare effect of foreign productivity and quality growth: A quantitative analysis.

Gawande, B. K., P. Krishna, and M. Olarreaga (2012). Lobbying competition over trade policy. *International Economic Review 53*(1), 115–132.

Gawande, K. and U. Bandyopadhyay (2000). Is protection for sale? evidence on the grossman-helpman theory of endogenous protection. *Review of Economics and statistics 82*(1), 139–152.

Gereffi, G. (1994). The Organization of Buyer-Driven Global Commodity Chains: How U.S. Retailers Shape Overseas Production Networks. In G. Gereffi and M. Korzeniewicz (Eds.), *Commodity Chains and Global Capitalism*, Chapter 5, pp. 95–122. Praeger Publishers.

Gereffi, G. (1999). International Trade and Industrial Upgrading in the Apparel Commodity Chain. *Journal of International Economics 48*(1), 37–70.

Gereffi, G., J. Humphery, and T. J. Sturgeon (2005). The Governance of Global Value Chains. *Review of International Political Economy 12*(1), 78–104.

Gereffi, G., S. Ponte, and G. Raj-Reichert (2019). *Handbook on Global Value Chains.* Edward Elgar Publishing.

Ghosh, A. (1958). Input-Output Approach to an Allocation System. *Economica 25*(97), 58–64.

Gil, R. (2011). The Interplay of Formal and Relational Contracts: Evidence from Movies. *The Journal of Law, Economics, and Organization 29*(3), 681–710.

Giri, R., K.-M. Yi, and H. Yilmazkuday (2021). Gains from trade: Does sectoral heterogeneity matter? *Journal of International Economics*, 103429.

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.

Goldberg, P. K. and G. Maggi (1999). Protection for Sale: An Empirical Investigation. *American Economic Review 89*(5), 1135–1155.

Gopinath, G. and B. Neiman (2014). Trade Adjustment and Productivity in Large Crises. *American Economic Review 104*(3), 793–831.

Gros, D. (1987). A note on the optimal tariff, retaliation and the welfare loss from tariff wars in a framework with intra-industry trade. *Journal of international Economics 23*(3-4), 357–367.

Grossman, G. M. and E. Helpman (1994). Protection for Sale. *American Economic Review 84*(4), 833–850.

Grossman, G. M. and E. Helpman (2002). Integration versus Outsourcing in Industry Equilibrium. *Quarterly Journal of Economics 117*(1), 85–120.

Grossman, G. M. and E. Helpman (2005). Outsourcing in a Global Economy. *Review of Economic Studies 72*(1), 135–159.

Grossman, G. M. and E. Helpman (2020). When Tariffs Disturb Global Supply Chains. NBER Working Paper No. 27722.

Grossman, G. M., E. Helpman, and A. Szeidl (2006). Optimal integration strategies for the multinational firm. *Journal of International Economics 70*(1), 216–238.

Grossman, G. M. and E. Rossi-Hansberg (2008). Trading Tasks: A Simple Theory of Offshoring. *American Economic Review 98*(5), 1978–1997.

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

Handley, K., F. Kamal, and R. Monarch (2020). Import Tariffs, Falling Export Growth: When Modern Supply Chains Meet Old-Style Protectionism. NBER Working Paper No. 26611.

Harms, P., O. Lorz, and D. Urban (2012). Offshoring along the Production Chain. *Canadian Journal of Economics 45*(1), 93–106.

Hausmann, R. and C. Hidalgo (2014). *The Atlas of Economic Complexity: Mapping Paths to Prosperity.* MIT Press.

Head, K. and T. Mayer (2019). Brands in motion: How frictions shape multinational production. *American Economic Review 109*(9), 3073–3124.

Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating trade flows: Trading partners and trading volumes. *The quarterly journal of economics 123*(2), 441–487.

Hertel, T. (2013). Global applied general equilibrium analysis using the gtap framework. In P. B. Dixon and D. W. Jorgenson (Eds.), *Handbook of Computable General Equilibrium Modeling*, Volume 1, pp. 815–876.

Hillberry, R. and D. Hummels (2013). Trade elasticity parameters for a computable general equilibrium model. In P. B. Dixon and D. W. Jorgenson (Eds.), *Handbook of Computable General Equilibrium Modeling*, Volume 1, pp. 1213–1269.

Hoang, T. (2020). The dynamics of global sourcing. mimeo Vanderbilt University.

Hofstede, G., G. J. Hofstede, and M. Minkov (2010). *Cultures and Organizations: Software of the Mind.* McGraw-Hill.

Huang, H. (2017). Germs, roads and trade: Theory and evidence on the value of diversification in global sourcing. mimeo London School of Economics.

Huang, Y., C. Lin, S. Liu, and H. Tang (2020). Trade Networks and Firm Value: Evidence from the US-China Trade War. mimeo.

Hummels, D., J. Ishii, and K.-M. Yi (2001). The Nature and Growth of Vertical Specialization in World Trade. *Journal of International Economics 54*(1), 75–96.

Hummels, D., D. Rapoport, and K.-M. Yi (1998). Vertical Specialization and the Changing Nature of World Trade. *Economic Policy Review 4*(2), 79–99.

Huneeus, F. (2018). Production network dynamics and the propagation of shocks.

Imbs, J. and L. Pauwels (2020). High Order Openness. CEPR Discussion Paper No. 14653.

IMF (2016). *Global Trade: What's Behind the Slowdown?* IMF World Economic Outlook.

Ito, T. and R. Baldwin (2021). The Smile Curve: Evolving Sources of Value Added in Manufacturing. *Canadian Journal of Economics*, forthcoming.

Ito, T., L. Rotunno, and P.-L. Vézina (2017). Heckscher–Ohlin: Evidence from Virtual Trade in Value Added. *Review of International Economics 25*(3), 427–446.

Jackson, M. O. (2010). *Social and economic networks.* Princeton University Press.

Jia, P. (2008). What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry. *Econometrica 76*(6), 1263–1316.

Johnson, R. C. (2018). Measuring Global Value Chains. *Annual Review of Economics 10*(1), 207–236.

Johnson, R. C. and A. Moxnes (2019). GVCs and Trade Elasticities with Multistage Production. NBER Working Paper No. 26018.

Johnson, R. C. and G. Noguera (2012). Accounting for Intermediates: Production Sharing and Trade in Value Added. *Journal of International Economics 86*(2), 224–236.

Johnson, R. C. and G. Noguera (2017). A Portrait of Trade in Value Added over Four Decades. *Review of Economics and Statistics 99*(5), 896–911.

Jones, L. P. (1976). The Measurement of Hirschmanian Linkages. *Quarterly Journal of Economics 90*(2), 323–333.

Jones, R. W. and J. P. Neary (1984). The positive theory of international trade. *Handbook of international economics 1*, 1–62.

Ju, J., H. Ma, Z. Wang, and X. Zhu (2020). Trade wars and industrial policy along the global value chains.

Kalemli-Özcan, S., S.-J. Kim, H. S. Shin, B. E. Sørensen, and S. Yeşiltaş (2014). Financial Shocks in Production Chains. mimeo.

Kamal, F. and H. Tang (2014). Relational contracts in global sourcing: Evidence from the US. mimeo John Hopkins SAIS.

Kasahara, H. and B. Lapham (2013). Productivity and the Decision to Import and Export: Theory and Evidence. *Journal of International Economics 89*(2), 297–316.

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.

Kashiwagi, Y., Y. Todo, and P. Matous (2021). Propagation of Economic Shocks through Global Supply Chains – Evidence from Hurricane Sandy. *Review of International Economics*, forthcoming.

Kee, H. L. and H. Tang (2016). Domestic Value Added in Exports: Theory and Firm Evidence from China. *American Economic Review 106*(6), 1402–1436.

Kehoe, T. J. (2005). An evaluation of the performance of applied general equilibrium models on the impact of nafta. In *Frontiers in Applied General Equilibrium Modeling: In Honor of Herbert Scarf*, pp. 378–401. Cambridge University Press.

Kehoe, T. J., P. S. Pujolas, and J. Rossbach (2017). Quantitative trade models: Developments and challenges. *Annual Review of Economics 9*, 295–325.

Keller, W. and S. R. Yeaple (2013). The gravity of knowledge. *American Economic Review 103*(4), 1414–44.

Kikuchi, T., K. Nishimura, and J. Stachurski (2018). Span of Control, Transaction Costs and the Structure of Production Chains. *Theoretical Economics 13*(2), 729–760.

Kim, S.-J. and H. S. Shin (2012). Sustaining Production Chains through Financial Linkages. *American Economic Review Papers and Proceedings 102*(3), 402–406.

Koopman, R., Z. Wang, and S.-J. Wei (2012). Estimating Domestic Content in Exports When Processing Trade is Pervasive. *Journal of Development Economics 99*(1), 178–189.

Koopman, R., Z. Wang, and S.-J. Wei (2014). Tracing Value-Added and Double Counting in Gross Exports. *American Economic Review 104*(2), 459–494.

Kramarz, F., J. Martin, and I. Mejean (2020). Volatility in the Small and in the Large: The Lack of Diversification in International Trade. *Journal of International Economics 122*, 103276.

Kremer, M. (1993). The o-ring theory of economic development. *The Quarterly Journal of Economics 108*(3), 551–575.

Krolikowski, P. M. and A. H. McCallum (2018). Goods-market frictions and international trade.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *American Economic Review 70*(5), 950–959.

Krugman, P. (1995). Increasing returns, imperfect competition and the positive theory of international trade. *Handbook of International Economics 3*, 1243–1277.

Kugler, M. and E. Verhoogen (2009). Plants and Imported Inputs: New Facts and an Interpretation. *American Economic Review Papers and Proceedings 99*(2), 501–507.

Kukharskyy, B. (2016). Relational contracts and global sourcing. *Journal of International Economics 101*, 123–147.

Laget, E., A. Osnago, N. Rocha, and M. Ruta (2020, September). Deep Trade Agreements and Global Value Chains. *Review of Industrial Organization 57*(2), 379–410.

Laugesen, A. (2018). Extensive Margins of Offshoring and Exporting. In W. Kohler and E. Yalcin (Eds.), *Developments in Global Sourcing*, Chapter 12, pp. 321. MIT Press.

Lee, E. and K.-M. Yi (2018). Global Value Chains and Inequality with Endogenous Labor Supply. *Journal of International Economics 115*, 223–241.

Lenoir, C., J. Martin, and I. Mejean (2019). Search frictions in international good markets.

Lenzen, M., D. Moran, K. Kanemoto, and A. Geschke (2013). Building EORA: A Global Multi-Region Input-Output Database at High Country and Sector Resolution. *Economic Research Systems 25*(1), 20–49.

Leontief, W. (1986). *Input-Output Economics* (second ed.). Oxford University Press.

Levchenko, A. A. and J. Zhang (2016). The evolution of comparative advantage: Measurement and welfare implications. *Journal of Monetary Economics 78*, 96–111.

Levin, J. (2003). Relational incentive contracts. *American Economic Review 93*(3), 835–857.

Li, X., X. Liu, and Y. Wang (2015). A Model of China's State Capitalism. mimeo.

Lim, K. (2018). Endogenous production networks and the business cycle. *Work. Pap.*

Liu, E. (2019). Industrial Policies in Production Networks. *Quarterly Journal of Economics 134*(4), 1883–1948.

Los, B., M. P. Timmer, and G. J. de Vries (2016). Tracing Value-Added and Double Counting in Gross Exports: Comment. *American Economic Review 106*(7), 1958–1966.

Los, B. and M. P. Timmers (2018). Measuring Bilateral Exports of Value Added: A Unified Framework. NBER Working Paper No. 24896.

Lu, D. (2010). Exceptional Exporter Performance? Evidence from Chinese Manufacturing Firms. mimeo University of Rochester.

Lu, W. (2019). Globalization, innovation, and firm dynamics. mimeo Xiamen University.

Macchiavello, R. and P. Miquel-Florensa (2018). Vertical Integration and Inter-Firm Relationships: Evidence from the Costa Rica Coffee Chain. LSE Mimeo.

Macchiavello, R. and A. Morjaria (2015). The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports. *American Economic Review 105*(9), 2911–2945.

Macchiavello, R. and A. Morjaria (2020). Competition and Relational Contracts in the Rwanda Coffee Chain. *Quarterly Journal of Economics*, forthcoming.

MacLeod, W. B. and J. M. Malcomson (1989). Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica: Journal of the Econometric Society*, 447–480.

Martin, J., I. Mejean, and M. Parenti (2020). Relationship Stickiness, International Trade, and Economic Uncertainty. Mimeo.

McCalman, P. (2004). Protection for sale and trade liberalization: An empirical investigation. *Review of International Economics 12*(1), 81–94.

McLaren, J. (2000). Globalization and vertical structure. *American Economic Review 90*(5), 1239–1254.

Melitz, M. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica 71*(6), 1695–1725.

Melitz, M. J. and S. J. Redding (2014a). Heterogeneous firms and trade. In G. Gopinath, E. Helpman, and K. Rogoff (Eds.), *Handbook of International Economics*, Volume 4, Chapter 1. Amsterdam: Elsevier.

Melitz, M. J. and S. J. Redding (2014b). Missing gains from trade? *American Economic Review 104*(5), 317–21.

Meng, B., Y. Zhang, and S. Inomata (2013). Compilation and Applications of Ide-Jetro's International Input-Output Tables. *Economic Research Systems 25*(1), 122–142.

Miller, R. E. and P. D. Blair (2009). *Input-Output Analysis: Foundations and Extensions* (second ed.). Cambridge University Press.

Miller, R. E. and U. Temurshoev (2017). Output Upstreamness and Input Downstreamness of Industries/Countries in World Production. *International Regional Science Review 40*(5), 443–475.

Miyauchi, Y. (2019). Matching and Agglomeration: Theory and Evidence from Japanese Firm-to-Firm Trade. mimeo Boston University.

Monarch, R. (2020). It's Not You, It's Me: Prices, Quality and Switching in US-China Trade Relationships. *Review of Economics and Statistics forthcoming*.

Monarch, R. and F. Kamal (2018). Identifying Foreign Suppliers in U.S. Import Data. *Review of International Economics 26*(1), 117–139.

Monarch, R. and T. Schmidt-Eisenlohr (2017). Learning and the value of trade relationships. *FRB International Finance Discussion Paper* (1218).

Morales, E., G. Sheu, and A. Zahler (2019). Extended gravity. *The Review of Economic Studies 86*(6), 2668–2712.

Morrow, P. M. and D. Trefler (2017). Endowments, skill-biased technology, and factor prices: A unified approach to trade. NBER Working Paper No. 24078.

Muûls, M. and M. Pisu (2009). Imports and Exports at the Level of the Frm: Evidence from Belgium. *World Economy 32*(5), 692–734.

Nagengast, A. J. and R. Stehrer (2016). Accounting for the Differences Between Gross and Value Added Trade Balances. *The World Economy 39*(9), 1276–1306.

Nunn, N. (2007). Relationship-Specificity, Incomplete Contracts and the Pattern of Trade. *Quarterly Journal of Economics 122*(2), 569–600.

Oberfield, E. (2018). A theory of input–output architecture. *Econometrica 86*(2), 559–589.

Olabisi, M. (2020). Input-Output Linkages and Sectoral Volatility. *Economica 87*(347), 713–746.

Orefice, G. and N. Rocha (2014). Deep integration and production networks: an empirical analysis. *The World Economy 37*(1), 106–136.

Ornelas, E. and J. L. Turner (2008). Trade liberalization, outsourcing, and the hold-up problem. *Journal of International Economics 74*(1), 225–241.

Ornelas, E. and J. L. Turner (2012). Protection and international sourcing. *The Economic Journal 122*(559), 26–63.

Ossa, R. (2011). A 'New Trade' Theory of GATT/WTO Negotiations. *Journal of Political Economy 119*(1), 122–152.

Ossa, R. (2015). Why trade matters after all. *Journal of International Economics 97*(2), 266–277.

Patel, N., Z. Wang, and S.-J. Wei (2019). Global Value Chains and Effective Exchange Rates at the Country-Sector Level. *Journal of Money, Credit and Banking 51*(S1), 7–42.

Puzzello, L. (2012). A Proportionality Assumption and Measurement Biases in the Factor Content of Trade. *Journal of International Economics 87*, 105–111.

Ramondo, N. and A. Rodríguez-Clare (2013). Trade, multinational production, and the gains from openness. *Journal of Political Economy 121*(2), 273–322.

Rauch, J. E. (1999). Networks versus Markets in International Trade. *Journal of International Economics 48*(1), 7–35.

Reimer, J. J. (2006). Global Production Sharing and Trade in the Services of Factors. *Journal of International Economics 68*, 384–408.

Reshef, A. and G. Santoni (2019). Are Your Labor Shares Set in Beijing? The View Through the Lens of Global Value Chains. CEPII working paper No. 2019-16.

Rodríguez-Clare, A., M. Ulate, and J. P. Vásquez (2020). New-keynesian trade: Understanding the employment and welfare effects of trade shocks. NBER Working Paper No. 27905.

Romalis, J. (2004). Factor Proportions and the Structure of Commodity Trade. *American Economic Review 94*(1), 67–97.

Ruffin, R. J. (1969). Tariffs, intermediate goods, and domestic protection. *The American Economic Review 59*(3), 261–269.

Rungi, A. and D. del Prete (2018). The Smile Curve at the Firm Level: Where Value is Added Along Supply Chains. *Economic Letters 164*, 38–42.

Sanyal, K. K. and R. W. Jones (1982). The theory of trade in middle products. *The American Economic Review 72*(1), 16–31.

Sforza, A. and M. Steininger (2020). Globalization in the time of covid-19.

Shapiro, J. S. (2020). The Environmental Bias of Trade Policy. *Quarterly Journal of Economics* (forthcoming).

Startz, M. (2016). The value of face-to-face: Search and contracting problems in nigerian trade. *Available at SSRN 3096685*.

Sugita, Y., K. Teshima, and E. Seira (2020). Assortative Matching of Exporters and Importers. mimeo.

Tang, H., F. Wang, and Z. Wang (2020). Domestic Segment of Global Value Chains in China under State Capitalism. *Journal of Comparative Economics*, forthcoming.

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.

Tintelnot, F. (2017). Global Production with Export Platforms. *Quarterly Journal of Economics 132*(1), 157–209.

Todo, Y., K. Nakajima, and P. Matous (2015). How Do Supply Chain Networks Affect the Resilience of Firms to Natural Disasters? Evidence from the Great East Japan Earthquake. *Journal of Regional Science 55*(2), 209–229.

Travis, W. P. (1964). *The Theory of Trade and Protection.* Number 121. Harvard University Press.

Trefler, D. and S. C. Zhu (2010). The Structure of Factor Content Predictions. *Journal of International Economics 82*, 195–207.

Tyazhelnikov, V. (2019). Production clustering and offshoring. *University of Sidney*.

Van Biesebroeck, J. and A. Schmitt (2020). Testing predictions on supplier governance from the global value chains literature.

Vandenbussche, H. and C. Viegelahn (2018). Input reallocation within multi-product firms. *Journal of International Economics 114*, 63–79.

Vanek, J. (1968). The Factor Proportions Theory: The N-Factor Case. *Kyklos 21*, 749–756.

Venables, A. J. (1987). Trade and trade policy with differentiated products: A chamberlinian-ricardian model. *The Economic Journal 97*(387), 700–717.

Vidal, C. J. and M. Goetschalckx (1997). Strategic production-distribution models: A critical review with emphasis on global supply chain models. *European Journal of Operational Research 98*(1), 1–18.

Walter, T. (2018). Trade and welfare effects of a potential free trade agreement between Japan and the United States.

Wang, J. (2021). Robot adoption, trade, and offshoring: A firm-level analysis.

Wang, X. and S. M. Disney (2016). The Bullwhip Effect: Progress, Trends and Directions. *European Journal of Operational Research 250*(3), 691–701.

Wang, Z., S.-J. Wei, X. Yu, and K. Zhu (2017). Characterizing Global Value Chains: Production Length and Upstreamness. NBER Working Paper No. 23261.

Wang, Z., S.-J. Wei, and K. Zhu (2013). Quantifying International Production Sharing at the Bilateral and Sector Levels. NBER Working Paper No. 19677.

Wicht, L. (2019). The impact of trade tensions on Switzerland: A quantitative assessment. *Aussenwirtschaft 70*(01), 1–34.

Wicht, L. (2020). The margin of importing sectors in the gains from trade.

World Bank (2020). *World Development Report: Trading for Development in the Age of Global Value Chains.* World Bank Publications.

Yang, C. (2020). Location choices of multi-plant oligopolists: Theory and evidence from the cement industry.

Yang, G. (2018). Infrastructure and global value chain position: Evidence from China.

Yeaple, S. R. (2003). The complex integration strategies of multinationals and cross country dependencies in the structure of foreign direct investment. *Journal of International Economics 60*(2), 293–314.

Yi, K.-M. (2003). Can Vertical Specialization Explain the Growth of World Trade? *Journal of Political Economy 111*(1), 52–102.

Yi, K.-M. (2010). Can Multi-Stage Production Explain the Home Bias in Trade? *American Economic Review 100*(1), 364–393.

Zhou, Y. (2020). The US-China trade war and global value chains.