Length Matters: The Value of a Longer Value Chain

Matthew Kidder*

This Rough Draft - March 2024: First Draft - November 2017

Abstract

This paper explores the role of network effects in international trade, focusing on value chain length as a novel metric within the Global Value Chain literature. Traditional trade theories, such as the Ricardian model, emphasize comparative advantage and industrial specialization, where countries allocate resources to their most efficient industries. However, this model overlooks the importance of complementarity in production, which generates gains from industrial variety and interconnected value chains. The paper contrasts Ricardian trade policy with industrial policies that promote variety and complex value chains.

This article finds that industrial network complexity is a source of comparative advantage in trade. Specifically, importers source from areas with more densely connected industries, a factor not explained by traditional sources of comparative advantage. The novel implication is that industrial specialization in relative comparative advantage can paradoxically lead to a trade disadvantage. The paper hypothesizes that the expansion of global value chains (GVCs) increases demand for skilled labor, especially in developed countries. Using an instrumental variables approach, the study finds that longer value chains cause skill effects, further supporting the idea that value chain length is a measure of technology.

JEL Classification: F13, F14, F43, F660, L520, P210, O250

*I thank Wang Zhi and Zhongxiu Zhao for founding the Research Institute for Global Value Chains and thus making this work possible. Their work provided access to the data and whose comments have directly benefited this work. We are also indebted to comments from Professors Heiwai Tang, Priya Ranjan, Bilal Khan as well as attendees at the Global Value Chain Workshop in Beijing. This paper was formerly titled: "Value Chain Length as a Source of Comparative Advantage, Varieties and Skill Upgrading Moderated by Levels of Development"

1 Introduction

Network effects in trade are not well understood. Much of trade theory, such as the Ricardian model, focuses squarely on specific industries as the means to understand trade patterns and potential gains from trade. Countries are prescribed free trade, which has the effect of inducing specialization within industries that have a relative comparative advantage. Industrial specialization is achieved through a Darwinian mechanism of survival of the fittest. Less productive domestic firms exit the market when forced to compete against more productive foreign firms. Industrial policy has long been used to protect domestic industry from foreign competition. And Industrial Policy has been under attack since Ricardian theory emerged in the wake of the Industrial Revolution. Ricardian theory argues that specialization leads to higher global production as countries allocate resources to industries in which they are most efficient.

Contrary to Ricardian specialization is the realization that complementarity in production creates gains to industrial varieties. Studying complementarity in production intermediates began as early as Leontief (1936). Jones (2011) builds a simple theoretical model that shows intermediate goods as a multiplier to total factor productivity. He uses this model to explain the great income discrepancies between countries. Dixit and Stiglitz (1977) build a model that shows varieties as a driver of economic growth. The Harvard Growth Lab bases their industrial policy on the principle of gains to variety and public sponsorship of industries close in the *know-how* space, Hidalgo Hausmann (2009). This type of industrial policy is opposite of Ricardian trade policy of comparative advantage, and this article brings these two camps together, by studying complexity within a context of comparative advantage.

Ricardian trade policy has a long geopolitical history dating back to post Industrial Revolution and reaching a height under the United States post World War II. The *Washington Consensus* pushed Ricardian Free Trade policies into the developing world. However, the global distribution of per-capita income has remained unchanged with few exceptions over the last 80 years.

On the other hand, China has emerged in defiance of free trade, by building industrial capacity through strategically positioned State Owned Enterprises. Often referred to as *Zombie Companies*, these zombies became landing points in adopting foreign technology, and disciplined Central Control allowed for many of these to die off or be spun into quasi-private firms. The effect of building industrial capacity is more interconnected domestic value chains. This article helps make sense of Chinese advantages in trade, which is not cheap labor.

This article will not tackle all topics but it will offer novel evidence on the significance of network effects in trade using value chain length as the central metric, and it will bring evidence to the skill effects from these connections. Our prior is that skill effects should be present if networks represent a kind of technology.

In this paper we explore one possible explanation for why demand for skills seems to be going up everywhere. Along with trade liberalization in developing economies has come the emergence of global value chains that break up the production process into many discrete steps that can be carried out in different countries. Most trade is now in intermediate goods. We think of global value chains developing in response to technological change as well as policy reform. Without big developing countries opening up, there would be less incentive to break up the production process; and without technological change, many of the GVC expansions would be impossible. Our hypothesis is that the expansion of global value chains has led to increased demand for skilled labor everywhere as skilled labor is needed for manage the GVCs and as inputs into important services that are complementary to value chains, such as finance, telecommunications, and transportation. If the factor-endowments effects are working as well, this could result in increased demand for skilled labor observed everywhere, but observed most strongly in developed countries, and less so in developing countries.

Instrumental variable identification for skill effects rests on a nice property of value chain length, which is that domestic value chain length transmits to longer global value chain lengths via trade channels; an importers global value chains become longer when an exporters domestic value chains become longer. The domestic value chain creates a foundation for global value chains, Koopman et al. (2015). This is a nice econometric property, since domestic policies that can influence value chain length in exporting countries can be exogenous to domestic economies in importing countries. This leads itself nicely to an instrumental variable approach.

But before getting too far ahead, we first need to describe in detail how value chain length is measured and why we think it might be important. Intuitive reasons, why length might matter, are given in Section 2. In Section 3, we carefully define how value chain length is measured and then we proceed to test value chain length as a source of comparative advantage. This section helps us to interpret value chain length as a measure of industrial inter-connectivity. Our untested prior is that denser industrial connections reduce cross-country transaction costs and also increase the amount of varieties. It may be counter-intuitive to think that more connection can reducing transaction costs. But consider the importers perspective, they take the existing industrial connectivity of each country as a given. The more connected the industries are within a country, the less the importer must intervene to form connections that might be needed to produce a given blue-print. Thus from the exporters' perspective, longer chains are transaction cost increasing, but from the importers' perspective, longer chains reduce transaction costs. The more interconnected industries are, the more access to intermediate goods, the more types of goods can potentially be produced without a need to form new supplier linkages, the less the transaction costs to foreign buyers, the more attractive the location becomes as a supply destination.

In Section 4 we examine the relationship between value chain length and skills, using OLS so that endogeneity issues are set aside for the moment. Value chain length is correlated with the use of high skill labor in developed economies, but not in developing ones. The medium skill labor share is positively correlated with value chain length in developing countries. The low-skill labor share is negatively correlated with value chain length in both high and medium income countries. The associations are in line with our hypotheses, but the causation could go either way.

The second half of Section 4 then uses an instrumental variables approach. We treat China's trade liberalization (tariff reductions) as an exogenous shock to value chain length. Kee and Tang (2015) show that Chinese import tariff reductions caused a restructuring within the Chinese economy, whereby Chinese firms substituted out of foreign intermediates and into domestic Chinese intermediates. The domestic content of exports actually rose as a result of falling trade costs because Chinese upstream intermediate goods manufacturers enjoyed costs savings from the import tariff reductions. We expand on Kee and Tang's findings and show that there is a clear link between domestic content and the value chain length. The process of substituting out of foreign intermediates and into domestic Chinese intermediates implies that the import tariff reductions promoted more domestic outsourcing, which has the effect of lengthening Chinese value chains. In turn, foreign buyers gain access to longer Chinese value chains, which lengthens the global value chains of China's downstream trading partners. In other words, a falling Chinese import tariff is a shock to other countries global value chain length. But this shock will vary with the inherent tradability of sectors, as well as the distance to different trading partners. By interacting the Chinese tariffs with measures of tradability and distance, we come up with an instrument that varies across sectors and trading partners. This instrument is a good predictor of trading partners' global value chain length. In the IV regressions upstream China GVC exposure has the following effect on the skill composition of the labor force: in high income countries, strong positive effect on high skills, moderate positive effect on medium skills, and a strong negative effect on low skills. In developing countries the results are modest positive effect on both high and medium skills, and a modesty negative effect on low skills. The results are consistent with the idea that the benefit of network effects is greater in denser network. Data and other technical topics are reserved for the Appendix.

2 Value Chain Length: Qualitatively Defined

For almost all final goods, production does not happen in one step. Instead, the production process involves a chain of producers that coordinate their production efforts and transform raw materials into final products. This is known as a value chain. Value chains are product level objects. A weakness of value chain research is that economy wide, value chain mapping is only done at the industry level.

Figure 1: Modules, Outsourcing and Value Chain Length

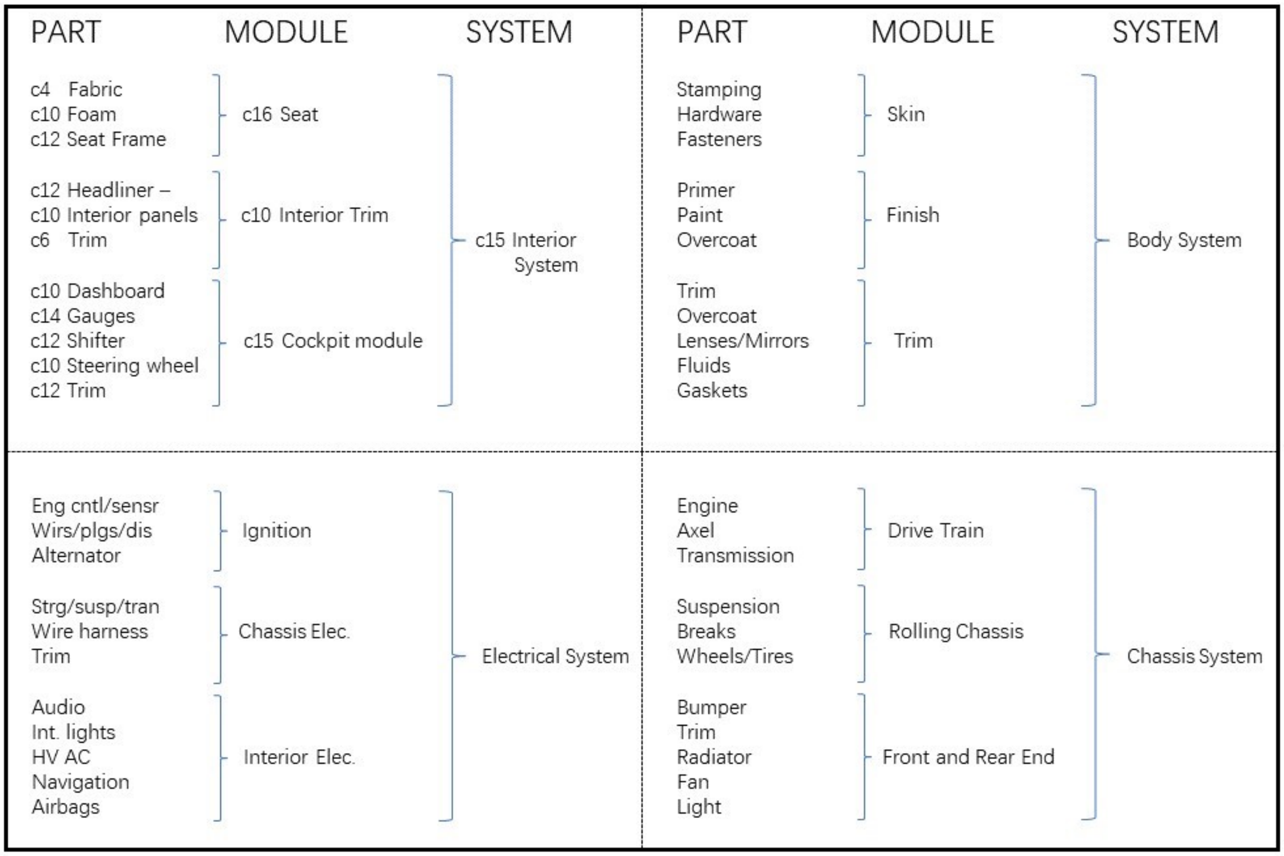


Table borrowed from Sturgeon and Lester (2003). Industry codes c1 - c34 are listed in the appendix. These were paired to the items in this graph by ad-hoc classification for demonstrative purposes. Actual industry codes for these products may be differently classified.

Industry level value chain measures are constructed using supply-use tables that map inputs from one industry into another. These supply-use tables form the "linkages" in the chains that is not available at the firm level, much less at the product level, for all firms or products within an economy. Product level measures from tear-down reports are available for individual products but these represent a small segment of products within an economy. Figure 1 is a partial tear-down report for automobiles.

Figure 1 is mostly borrowed from Sturgeon and Lester (2003). They found an increasing preference for modulization in auto production. For example, vehicle doors can be delivered with glass, fabric and mirrors pre-assembled. Parts create modules, which are combined into systems, which are then processed into a final product, a car. They find that modules are natural points for outsourcing and that 75 percent of a vehicle can be accounted for by 15 modules.

We modified Figure 1 by adding ad-hoc 2-digit WIOD industrial classifications to the Interior System panel. Doing so highlights several points. First, industrial classification is largely determined by the industrial organization. For discussion, define P Parts, M Modules, S Systems that make an F

Final product. The entire Figure reduces to a single classification, industry "c15", when one company manufactures all parts of production in-house, when the chain is F. We know that such organization is unrealistic so let us consider the extreme case of vertical integration to be $P \to F$. Note that this form of industrial organization has the highest number supplier relationships, from the perspective of the final product producer. There are a total of 44 parts in Figure 1, so assuming each part is produced by one company this corresponds with 44 relationships.

Now consider what happens when all modules are outsourced. The industrial organization becomes $P \to M \to F$. Several changes are important to note. The value chain length became longer because of outsourcing. The number of suppliers that the final good producer interacts with fell from 44 to 12. If we abstract, and assume this company represents an entire economy, then the number of varieties that are available also increased, with the addition of module goods that were previously less available. But, the final product has not changed, nor the technology requirement needed to produce it.

The final product did not change because we assume this is fixed by blue-print design. In other words, the parts, modules and systems are all perfectly complementary in the final product production. Something else interesting happens; let us forget this company as representing the economy and allow other final goods to exist besides cars. When there is modularization then there is scope for modular substitutability between final products. For example, a seat company can supply to Interior Systems for cars but they can also supply seats to buses, trains, airplanes or compete in household furniture industries. Seat technology can be substituted between product but not within product; a seat cannot substitute for an engine. Having a seat module that is available in aircraft production has a similar effect of reducing the number of suppliers that the aircraft maker needs to interact with in order to product a final product. Substitution between products leaves the door open for potential externalities and market failures, which will not be developed in this article but are interesting to consider.

3 Value Chain Length - Measurement and Evidence

Our primary data source for trade is the World Input Output Database, and our value chain measures are borrowed from the Research Institute for Global Value Chains at UIBE in Beijing, which calculate value chain length proposed by Antras and Chor (2017). The value chain measures are constructed from the WIOD, thus our first task it to outline the construction of our key measurements. Table 1 illustrates the data structure of the WIOD and also serves to define the variables in our models.

The input-output like structure of the WIOD data, in Table 1, describes a mapping of value addition between all industry-country pairs. We have G countries an N industries in this world economy. The Z matrix represents intermediate good flows and each Z^{rs} element in Table 1 are each an $N \times N$ matrix of inter-industry intermediate flows with elements z_{ji}^{sr}. The element z_{ji}^{sr} is defined

Table 1: WIOD Variables

Outputs / Inputs		Intermediate Use				Final Demand				Total Output
		1	2	$\cdots$	G	1	2	$\cdots$	G	
Intermediate Inputs	1	Z^{11}	Z^{12}	$\cdots$	Z^{1g}	Y^{11}	Y^{12}	$\cdots$	Y^{1g}	X^1
	2	Z^{21}	Z^{22}	$\cdots$	Z^{2g}	Y^{21}	Y^{22}	$\cdots$	Y^{2g}	X^2
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	G	Z^{g1}	Z^{g2}	$\cdots$	Z^{gg}	Y^{g1}	Y^{g2}	$\cdots$	Y^{gg}	X^g
Value-added		Va^1	Va^2	$\cdots$	Va^g					
Total input		$(X^1)'$	$(X^2)'$	$\cdots$	$(X^g)'$					

Notes: This table is borrowed from Wang et al (2017)

as the intermediate good flow of industry j, source country s in production of industry i, destination country r production.

$$Z^{sr} = \begin{vmatrix} z_{11}^{sr} & z_{12}^{sr} & \cdots & z_{1N}^{sr} \\ z_{21}^{sr} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ z_{N1}^{sr} & \cdots & \cdots & z_{NN}^{sr} \end{vmatrix}$$

Total output X^s in Table 1 is an $NG \times 1$ column vector, represented as X, with individual elements, x_j^s. X is calculated by summing across the columns of intermediate goods, Z and final goods Y in Table 1; $X = Z + Y$. The input-output insight is that intermediate flows Z are themselves derived from total output. An A fraction of total output X is used to create Z intermediate flows according to Z=$A\hat{X}$. We introduce the hat notation on the column vector X, to denote that the column vector is diagonalized over a square matrix.

The matrix A has dimension $NG \times NG$ and is critically important in our future analysis. The A coefficient matrix is what governs the creation of intermediate flows; $z_{ji}^{sr} = a_{ji}^{sr} x_j^s$. The a_{ji}^{sr} element maps all of the links between industries and between countries. The A matrix can be grouped into country blocks similarly to how the Z matrix is partitioned in Table 1. For example, A^{sr} describes how X^s total output in sourcing country s is used in the creation of intermediate goods in importing country r. When $s = r$ then A^{sr} shows how total output from s is mapped into the creation of intermediate domestic industries withing s; this is a domestic A coefficient matrix. Our research question, seeks to understand the domestic characteristics of exporters that determine sourcing location of importers.

Thus we will construct our independent variable using the domestic A matrix. This amounts to us looking at A coefficients that are embedded down the diagonal country blocks of the A matrix, which corresponds to the diagonal country blocks of the Z matrix.

Getting back to Table 1, we can define $Z = A\hat{X}$ and express the summation across the columns $X = Z + Y$ as:

$$X = A\hat{X} + Y \tag{1}$$

Equation 1 moves along the rows, across the columns of Table 1 and thus describes where inputs are used in production. A nice feature about Table 1 is that we can also see where a country-industry sources it's intermediates by moving down the column, across rows. Summing down the column, across the rows, derives another accounting identity, $X' = Z' + Va'$, where X' is the total input rather than output. This vector is transposed into a $NG \times 1$ column vector. In this model, a B fraction of inputs are transformed back into inputs, such that $Z = \hat{X}'B$.

$$X' = \hat{X}'B + Va' \tag{2}$$

The interpretation of the B matrix is slightly different as it transforms inputs into inputs under this construction. It is natural to think of Equation 2 to be upstream looking, whereas Equation 1 looks down the production stream. Our primary research question asks, what factors in the exporting countries, might influence the intensive margin sourcing in importing countries. We are interested in sourcing, which makes Equation 2 a more natural fit for our research question than Equation 1.

In summary, answering our research question depends on us first deriving a measure of sourcing preference and then deriving a measure that characterizes the connectivity in the domestic blocks of B. International sourcing preferences should naturally be calculated up the columns, excluding the domestic segment, and then compared with a measure of the domestic B matrix for each country where $r \neq s$. Table 1 gives an example where data for our dependent (blue highlighted blocks) and independent variable (red highlighted blocks) come from if the country = 1. The next two subsections will now describe how these measures are constructed in more detail.

Measurement of Dependent Variable: As a first step, we need to clearly define our dependent variable of interest, which is a type of weighted import share. Define the import share, s_{ji}^{sr} as the intermediate good imports from (s, j) divided by total intermediate good imports in (i, r) production:

$$s_{ji}^{sr} = \left(\frac{z_{ji}^{sr}}{\sum\limits_{s \forall s \neq b} \sum\limits_{j} z_{ji}^{sr}} \right) \tag{3}$$

The denominator of s_{ji}^{sr} varies across importing industries i and importing countries r. The effect is to normalize across industries, within country, to reflect that some industries are more import oriented than others. This is a type of interaction between country-industry fixed effects on the demand side. The equation looks similar to the dependent variable in Levchenko (2007), with a few key differences. The biggest difference is that we are using bilateral data at the country-industry level, whereas Levchenko's data does not have importing industries. Importing industries are important to us because we are interested in value chains, which are specific to an industry.

We cleanse scale effects and trade closeness similar to Levchenko by dividing s_{ji}^{sr} by the average (j,s) shares in all industries within country r. The main difference is that we have multiple importing countries, whereas Levchenko only studied US imports.

$$share_{ji}^{sr} = \left(\frac{s_{ji}^{sr}}{\frac{1}{N_i N_j} \sum_{i}^{N_i} \sum_{j}^{N_j} s_{ji}^{sr}} \right) \tag{4}$$

This weighted share is our main object of interest because it measures the location choice of supply on an intensive margin, net scale effects. When this measure is high for a given industry, country (j,s) then it reveals that industry-country (i,r) has a preference for sourcing from (j,s). We are interested in understanding what drives these sourcing destination decisions and so our analysis will examine characteristics $X_{j,s}$, in the sourcing industry-country pair (j,s) that relate to $share_{ji}^{sr}$. Thus our econometric model will have $share_{ji}^{sr}$ as the dependent variable, with various types of characteristics, $X_{j,s}$, as the explanatory variables.

Measurement of Value Chain Length: The length of a value chain is a measurement that characterizes how intensely intermediates flow within an economy. The length can be derived directly from iterating on Equations 1 or Equation 2 and counting steps at each iteration, which we will call the "length" that is associated with that step. Our measure of value chain length is that presented by Antras and Chor (2017) and Wang et al (2017). The measure is created as follows: If we iterate on Equation 2 and count each iteration as a value chain step, then we can derive an upstream measure of value chain length. As we will see, the main factor that determines the length is the fullness of the Z matrix, which determines the fullness of the B matrix of coefficients.

Equation 5 illustrates how value chain length is counted. In the first stage, some value added is made directly using primary inputs according to Equation 2. These have only one step in the value chain, which we notate with a (1). B fraction of X is used to produce intermediate flows which enter a second stage of production, as seen in the second line of Equation 5. Value added in this second stage is produced with inputs that underwent two steps in the value chain and we notate this with a (2) in Equation 5. Continuing this process, we arrive at the N^{th} stage that has an associate value chain length of $(N+1)$.

$$\begin{aligned}
&1^{st} \text{ Stage } L^u = XB + (1)Va' \\
&2^{nd} \text{ Stage } L^u = (XB + Va')B + Va' = XB^2 + (2)Va'B + (1)Va' \\
&3^{nd} \text{ Stage } L^u = (XB + Va')B^2 + Va'B + Va' = XB^3 + (3)Va'B^2 + (2)Va'B + (1)Va' \\
&\qquad\qquad \vdots \\
&N^{th} \text{ Stage } L^u = XB^n + Va'\left(\sum_{n=0}^{N}(N+1)B^n\right)
\end{aligned} \tag{5}$$

Let S denote the sum as N tends towards infinity. XB^n is converging to zero in the limit because the elements in B are less than unity. It is straight forward to show that $(S - SB) = Va'(I + B + B^2 + B^3 + ...) = Va'(I - B)^{-1}$. Define $W = (I - B)^{-1}$. Thus $S = Va'WW$. Antras and Chor (2017) suggest to weight Equation 5 by gross inputs X. Thus the upstream length, according to Antras and Chor is:

$$L^u = Va'WW\hat{X}^{-1} \tag{6}$$

We borrow this measure, which is directly made available by the Research Institute for Global Value Chains at the University of International Business and Economics in Beijing. The documentation for this data set is in Wang et al (2017). Equating the documentation in Wang et al (2017) with Antras and Chor (2017) involves a bit of matrix manipulations. First note that directly iterating on Equation 2 produces:

$$X = Va'(I - B)^{-1} = Va'(I - \hat{X}^{-1}Z) = Va'W$$

This reduces Equation 6 into $L^u = \mathbf{1}\hat{X}W\hat{X}^{-1}$. $W = (I - B)^{-1}$ can be related to the Leontief inverse matrix, which we have yet to define. The Leontief inverse matrix can be obtained if we iterate on Equation 1:

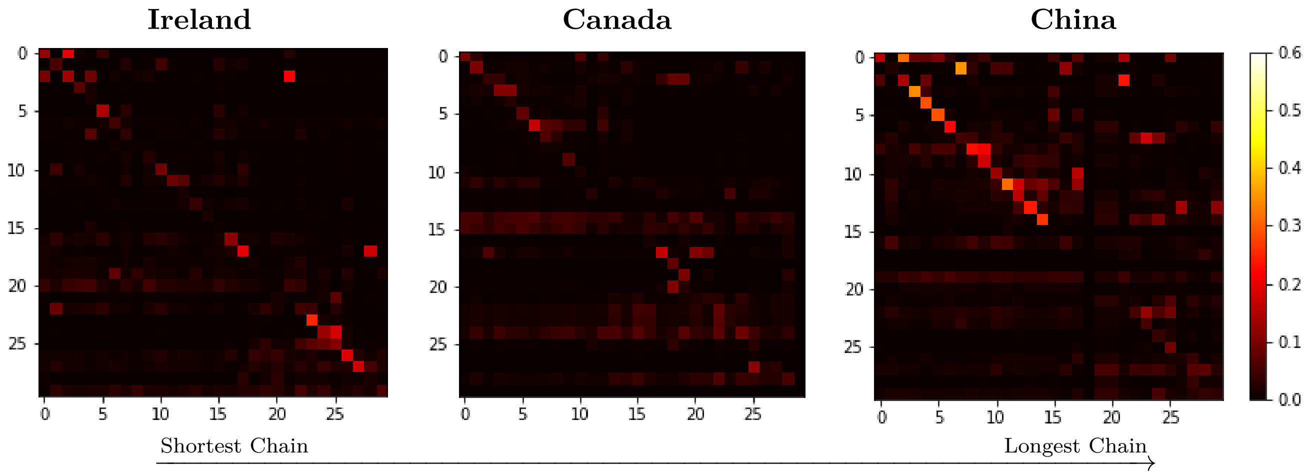
$$X = Y(I - A)^{-1} = Y(I - Z\hat{X}^{-1})^{-1} = YL$$

Where $L = (1 - A)^{-1}$ is the Leontief inverse matrix. It can be shown that $W = \hat{X}^{-1}L\hat{X}$, thus reducing Equation 6 to the equivalent upstream value chain length measure that is in Wang et al (2017).

$$L^u = \mathbf{1}L \tag{7}$$

Equation 7 is nice because it is compact and easy to calculate using data. All we need to do is calculate the Leontief inverse matrix and then sum across the rows, up the columns.

Table 2: Coefficient Matrix A



Notes: These heat maps plot the domestic A matrix for each country. Industries are on the axis and the color corresponds to the scale with a minimum=0 and a maximum=0.6

The domestic upstream value chain length is a measure of how interconnected industries are within a country. Value chain length is a measure of the density of the A matrix. The A matrix, or closely related B matrix is what drives the discounting of longer segments of the value chain, as we saw in Equation 5. Matrices with larger coefficients and more non-zero coefficients result in longer value chains. The coefficient matrix dominates the value chain length metric because these are sparse matrices and thus discounting is rapid. Figure 2 plots the A matrix as a heat map for Ireland, Canada and China. These countries are ordered from shortest to longest in terms of value chain length, with China having the longest value chains in the world. Each number on the axis represents an industry in the WIOD[1]. One striking feature is the amount of dark space in these matrices, showing that they are sparse. The A matrix of China has less dark space than Ireland and also has stronger coefficients. The interpretation is that the linkages between industries, as expressed in terms of intermediate good flows, is stronger in China versus in Ireland. We will formally use upstream value chain length, as calculated by Equation 6 and interpret our findings as a measure of industrial connectivity.

Main Findings in Graphical Form: The positive correlation between $share_{ji}^{sr}$ and value chain length in source country s industry j is a very robust feature in the data and can be directly observed in Figure 2. Figure 2 shows the sourcing countries in which, industry 16 sources its machinery (industry 13). Each plot represents a single country's production of industry 16. For compactness, this is plotted for only four countries, Malta, China, Germany and the United States. Each point on the plot

[1]Industry numbers and industry names are listed in the Appendix

represents the country that supplies machinery as an intermediate input.

Figure 2: **Sourcing Destination and Value Chain Length**
Year=2008, Industry=16

Machinery Industry 13 used in Manufacturing Industry 16 in Four Countries

On the x-axis is the domestic value chain length of the sourcing country[2]. On the y-axis, we have $share_{ji}^{sr}$. The plots show a positive correlation between the value chain length of the trading partner and the intensity of sourcing from that partner. In other words, there seems to be a sourcing preference for value chain length in ones suppliers. This is a novel finding. But this positive relationship can be confounded by other sources of comparative advantage, which is why we will now formalize this preliminary plot within a more robust regression framework that allows us to control for likely confounders.

[2]The value chain length is taken from the Research Institute for Global Value Chains measures, which are publicly available. The measures of value chain length are derived by Zhi et al. (2017), whose main contribution is decomposing value chain length into a strictly domestic component and into other trade groups based on value added. This allows us to measure the domestic value chain length separately from the value chain length of trade with upstream countries.

3.1 Statistical Model of Comparative Advantage and Value Chain Length

Using bilateral trade data, we closely follow Levchenko's (2007) empirical model for testing sources of comparative advantage. The dependent variable is $share_{ji}^{sr}$ and is a general measure of sourcing preference. Sourcing preferences naturally model sources of comparative advantage with models that take a form such as:

$$share_{ji}^{sr} = \beta_0 + \beta_1 (\text{industry characteristics})_j \times (\text{country endowments})_s + \alpha_j + \alpha_s + \alpha_i + \alpha_b + \epsilon_{sr}$$

The common analysis recognizes that industries are different in terms of their production technologies and contracting intensities. Some industries, such as steel refinement are more capital intensive than apple farming. Countries have different endowments of capital, thus the interaction of country level capital endowments and industrial level capital intensity, as an independent variable, test is there is a sourcing preference for capital intensive goods from countries with higher capital endowments. The key parameter of interest is thus β_1. In another example, country level institutional quality, such as rule-of-law, can reduce transaction costs in industries that are more contract intensive, Ranjan and Lee (2007), Nunn (2007) and Levchenko (2007). The significance of the interactions are made more dominant by introducing country and industry fixed effects, which absorb level effects.

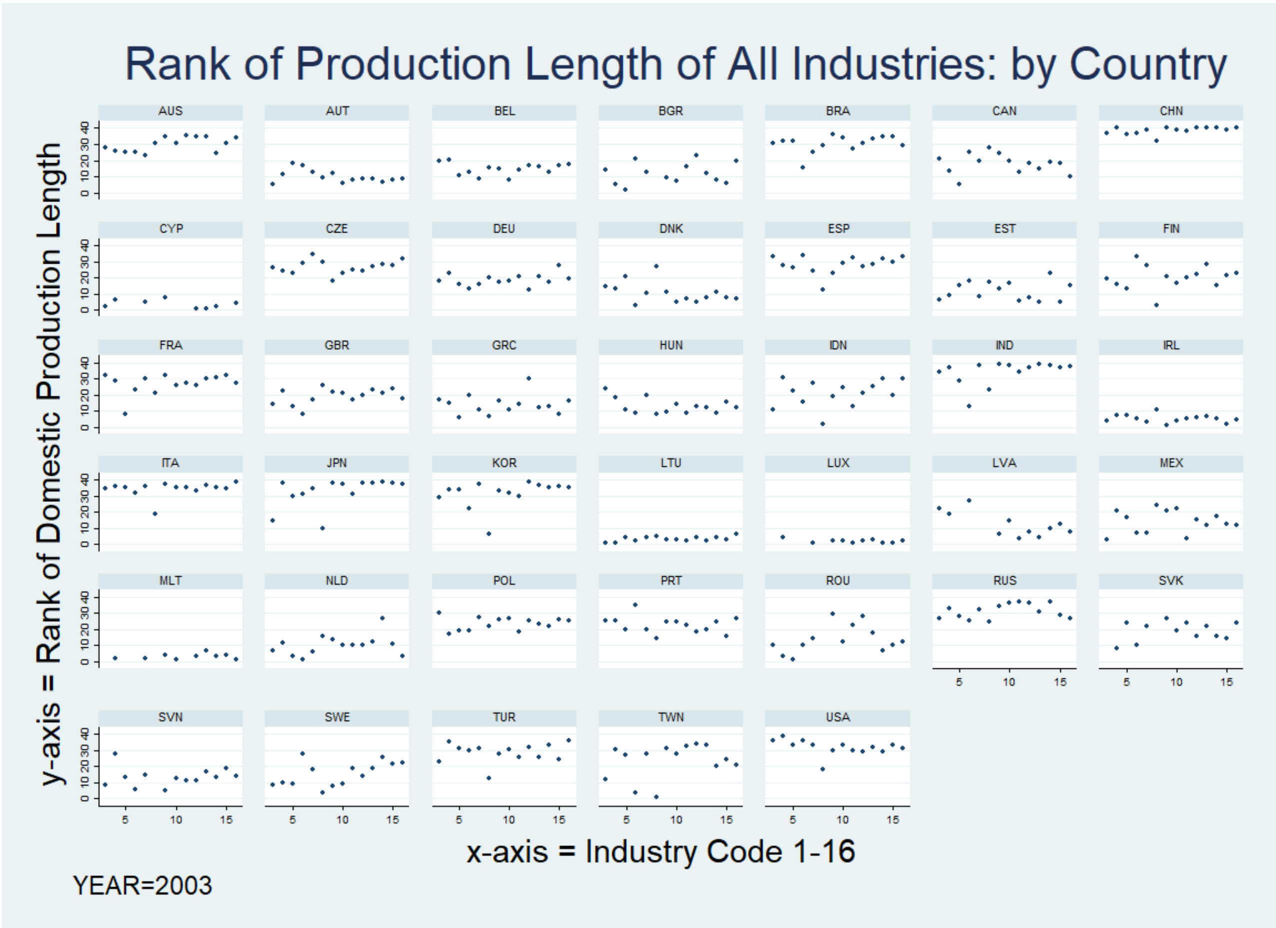
This is the model that we will use to study how a suppliers value chain length is related to sourcing preference for that supplier. In this case, we are also interested in the variation between industry and country level interactions. And this brings up a critical question; is value chain length a country characteristic or an industry characteristic or both?

Understanding the variation of value chain length is critical for interpreting the coefficients on value chain length within our model. Firstly, value chain length varies by industry, with some industries having longer chains than others as a result of their inherent product complexity. The chain length for agriculture products should be expected to be much shorter than the value chain for consumer electronics, for example. Secondly, we find that value chain length is also a type of country endowment, whereby some countries have longer chains for all industries. In other words, there are country level effects in value chain length. Thirdly, variation in value chains are driven by an interaction between country and industry effects. For example, some industries are more interlinked, more developed, and thus have longer chains within a country.

In fact, we find that when we rank each country based on value chain length, within each industry, and then plot the rank orders over industries, then we see clearly defined level effects. This is exactly what we do in Figure 3.

In Figure 3, the x-axis is the industry number and the y-axis is the country rank of value chain

Figure 3: **Rank of Production Length: by Country-Industry**



Countries are ranked for each industry, with 40 being associated with the longest production chains in that industry and 1 being the shortest. The plots thus represent the ranking of all industries within each country.

length within that industry. A rank of 40 is the highest rank, which can be interpreted as that country having the longest value chain, relative to other countries, for that given industry. There is a clear country level effect. China has the longest value chains in the world for most industries, highlighting a novel strength of the Chinese manufacturing complex. Further, there is within country variation across industries but this is not caused by differences in length between industries since rankings are within industry. The last interesting observation, is that countries with the shortest chains have no industries with longer chains, which is a striking feature since the opposite is not always true; some countries with long chains can have one or two industries with relatively shorter chains. It seems that there are barriers in terms of value chain length. In summary, there are clear country level variations in value chain length and also variation in length within country. In order to isolate comparative advantage, we will thus specify a model in which variation is across countries, within an industry, and value chain length will enter without any interaction in our baseline model.

Table 3: Value Chain Length and Comparative Advantage: Baseline Model

dep var: $share_{ji}^{sr}$				
	(1)	(2)	(3)	(4)
Supplier Production Length	1.477***	1.192***	1.473***	
	(0.036)	(0.040)	(0.036)	
(skill intensity)×(skill endowment)	13.470***	12.416***	12.824***	13.260***
	(0.447)	(0.451)	(0.451)	(0.452)
(capital intensity)×(capital endowment)	0.060	0.193***	0.171***	0.266***
	(0.042)	(0.042)	(0.043)	(0.043)
Supplier Herfindahl		-1.769***		
		(0.106)		
Inst. × Herfindahl (benchmark)			7.816***	8.081***
			(0.661)	(0.663)
Constant	-8.895***	-7.361***	-15.373***	-12.726***
	(0.246)	(0.262)	(0.600)	(0.599)
Country-Industry FE	YES	YES	YES	YES
Observations	193916	193916	193916	193916
Adjusted R^2	0.419	0.420	0.420	0.415

Notes: institutional quality is measured as "rule of law" from Kaufmann, Kraay, and Mastruzzi (2003). Capital and Labor endowments are taken from Hall and Jones (1999) and are in natural logs. Chain Length is the domestic upstream length and is taken from the Research Institute of Global Value Chains, see Wang et al (2017). The Herfindahl Index, skill and capital intensities are calculated from World Input-Output Data; see Timmer et.al. (2015). Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Our baseline model is as follows. We notate sourcing countries with either subscript or superscript s and countries who buy goods with b, reflecting the bilateral nature of our data. Industries are notated with either i or j. The dependent variable follows Levchenko (2007) and is the normalized import share. Thus the dependent variable $share_{ji}^{sr}$ is the import share of industry j from source country s in the production of industry i country b; normalized by dividing by the average import share of source industry j country s in other industries within country b. Normalizing in this way has the effect of accounting for the closeness of trading relationships as well as for economic size and thus makes coefficients more comparable across countries. Our baseline model is:

$$share_{ji}^{sr} = \beta_0 + \beta_1 V_{sr} + \beta_2 hhi_j \times inst_s + \beta_3 skint_j \times skill_s + \beta_4 capint_j \times cap_s + \alpha_j + \alpha_b + \alpha_s + \alpha_i + \epsilon_{ji}^{sr} \quad (8)$$

Time scripts are omitted since we perform this analysis in a cross section of industries in the year 1998. The domestic value chain length of supplier in country s industry j is V_{sr}. We also include skill intensity of industry j, denoted $skint_j$ and capital intensity, denoted $capint_j$[3]. Each of these

[3]These estimates are benchmarked to the US.

Table 4: Value Chain Length, Institutions and Comparative Advantage

dep var: $share_{ji}^{sr}$	(1)	(2)	(3)	(4)
Supplier Chain Length	1.799***	1.381***	1.436***	1.155***
	(0.140)	(0.142)	(0.144)	(0.145)
Supplier Chain Length × Inst. Quality	-0.441**	-0.257	0.050	0.090
	(0.186)	(0.186)	(0.190)	(0.190)
(skill intensity)×(skill endowment)	13.349***	12.351***	12.834***	12.071***
	(0.450)	(0.454)	(0.452)	(0.455)
(capital intensity)×(capital endowment)	0.057	0.190***	0.172***	0.262***
	(0.042)	(0.042)	(0.043)	(0.043)
Supplier Herfindahl		-1.760***		-1.578***
		(0.106)		(0.108)
Benchmarked Herfindahl × Inst. Quality			7.855***	5.851***
			(0.678)	(0.691)
Constant	-8.704***	-7.258***	-15.427***	-12.415***
	(0.258)	(0.273)	(0.635)	(0.667)
Country-Industry FE	YES	YES	YES	YES
Observations	193916	193916	193916	193916
Adjusted R^2	0.419	0.420	0.420	0.420

Notes: institutional quality is measured as "rule of law" from Kaufmann, Kraay, and Mastruzzi (2003). Capital and Labor endowments are taken from Hall and Jones (1999) and are in natural logs. Chain Length is the domestic upstream length and is taken from the Research Institute of Global Value Chains, see Wang et al (2017). The Herfindahl Index, skill and capital intensities are calculated from World Input-Output Data; see Timmer et.al. (2015). Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

in interacted with country level skill and capital endowments, $skill_s$ and cap_s respectively, which are borrowed from Hall and Jones (1999). Including these helps further control potential sources of spurious correlations driven by differences in returns to capital and labor across countries. We also include country and industry fixed effects to add additional controls for factors like institutional quality and industrial differences like tradability. Column 1 in Table 3 reports the OLS estimates of the naive baseline model. We see that value chain length is positively and significantly related to import shares, which is consistent with value chain length being a source of comparative advantage.

Next, we include the suppliers Herfindahl Index of intermediate goods to eliminate the possibility that the coefficient on value chain length is simply being driven by the first upstream stage. The Herfindahl Index of intermediate goods concentration measures product complexity one step upstream while production length accounts for complexity along the entire chain. With the Herfindahl index included, the coefficient on value chain length in Column 2, now reflects deeper supplier relationships. The coefficient on value chain length is only marginally smaller than before and still significant. The results are robust to the Herfindahl index being benchmarked to the US, as a type of industrial fixed effect or not benchmarked, thus allowing the index to have full variation across countries and industries[4].

But institutions matter in determining transaction costs and thus can influence industrial organization. Thus, as a first step at controlling for the possibility of institutional quality, we include the Herfindahl Index interacted with country level institutional quality, denoted $inst_s$, which is taken from Nunn (2007). Column 3 report shows that the value chain length is robust to a basic accounting of institutional quality as a source of comparative advantage and Column 4 is included to make our results comparable to those of other authors.

But we may still be concerned that institutional effects are not being completely purged from prior specifications. Transaction costs are perhaps even more relevant in long value chains than they are one step upstream. Table 4 addresses this possibility by including interacting upstream value chain length with institutional quality. Institutional interactions with value chain length are negative and significant only when the Herfindahl is not included. There is only weak evidence to support a transaction cost theory, that longer value chains reduce transaction costs. The coefficient on the interaction between value chain length and institutional quality is mostly small and insignificant when interacted with length, while institutions remain significant when interacted with the Herfindahl Index. The coefficient on value chain length remains very robust across all of these specifications, both in terms of size and significance.

In summary, value chain length is a robust source of comparative advantage. The source of advantage comes from industrial organization that is deeper than one step upstream. There are at least two reasons why longer value chains might matter. First, they can reduce transaction costs to foreign buyers; the empirical evidence is weak on this point. A second possibility is that longer value chains create more varieties within an economy. This theory has not been directly tested in this article. What we have shown is that longer value chains do matter, even after controlling for other sources of comparative advantage. Our finding is significant in itself. Future work can be done to determine how value chains become longer and why they might matter. For now, we are resolute that value chain length is important and a robust source of comparative advantage.

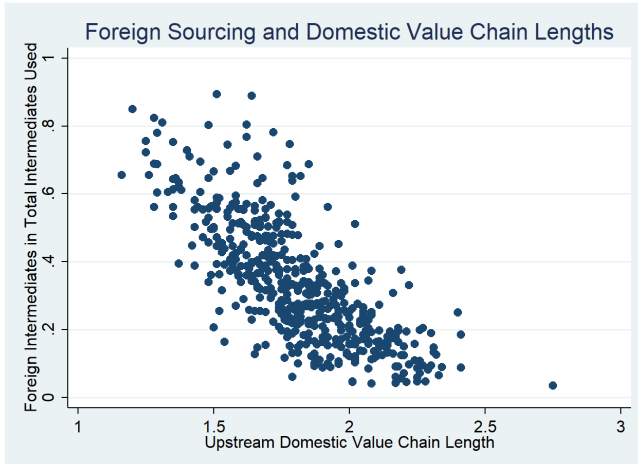
3.1.1 Domestic Value Addition and Policy Implications

Figure 4 shows that longer domestic value chains are negatively/positively correlated with foreign/domestic intermediate shares. This simple plot directly links our work with Kee and Tang (2015). They noted a surprising result that falling Chinese import trade barriers resulted in a deepening of Chinese domestic content. Figure 4 shows a positive relationship between domestic shares of intermediates and domestic value chain length. In Figure 3, China clearly has always had long value chains relative to

[4]The benchmarked Herfindahl Index of intermediate goods, denoted hhi_j strictly follows Nunn (2007) and Levchenko (2007)

Figure 4: **Foreign Sourcing and Domestic Value Chain Length**

Year=2000

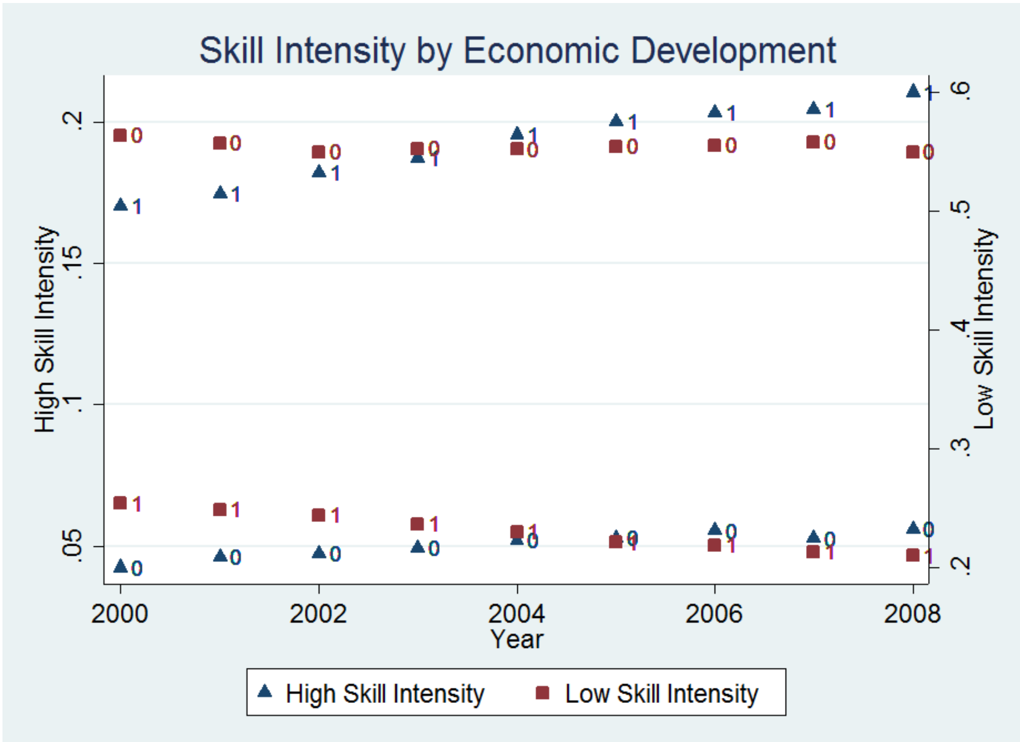


other countries. In other words, China's longer value chains existed prior to China's entry into the WTO. Thus the deepening of domestic content that Kee and Tang noted was made possible by the depth of the Chinese production network. We argue that it is longer chains that make substitution into domestic components possible. China has been well known for its ability to copy goods. There are stories of copied goods being sold in the market before actual goods are formally released. This capacity reveals a strength of the Chinese economy, to absorb foreign technologies. China has the longest value chains in the world and perhaps this is a valuable lesson of industrial organization that the rest of the developing world can gain from.

4 GVC Length and Distribution of Skills

Based on the last section, we found that value chain length can be a source of comparative advantage. This implies that lengthening value chains should have an effect on factor markets through a trade channel. Factor market reallocation that is induced by trade should follow production factor sources of comparative advantage, with increased flows of lower skill intensive production into lower skill abundant countries. At the same time, the industrial reorganiation from shorter value chains to longer value chains can require higher skills as longer value chains place a heavier burden on supply chain management and quality controls than shorter chains. In this way, longer value chains can have an upskilling effect in all countries including those that are abundant in lower skills. For these lower skill abundant countries, the two effects move in opposite directions, thus determining which dominates is an empirical question. Figure 5 plots weighted average skill intensity by level of economic development between 2000 to 2008. This plot shows that both the developing and developed world

Figure 5: **Skills in High Income and Medium Income Countries**



High income countries are marked with a "1" and Medium income countries are marked with a "0"

experienced skill upgrading during this period[5]. The purpose of this section is to determine to what extent lengthening of value chains might be capable to explain these patterns in the distribution of skills.

This section begins with an ordinary least squares approach to evaluate the correlation between value chain length and the distribution of skills. Fixed effects estimates are reported in the Appendix. We conclude this section with an instrumental variables approach. All of our methods produce qualitatively consistent results and so we will focus on the OLS model to build intuition. The correlation of skills and value chain length can be evaluated using country-industry-year level data from the World Input Output Database, within the following ordinary least squares regression model:

$$L^s_{ijt} = \beta_0 + \beta_v V_{ijt} + \beta_d D_j + \beta_{dv}(V_{ijt} \times D_j) + \boldsymbol{\beta_x X_{ijt}} + \boldsymbol{\beta_{dx}(X_{ijt}} \times D_j) + \alpha_i + \alpha_t + \epsilon_{ijt} \tag{9}$$

We abandon subscript notation from the last section and re-define it here. On the left is the share of skill s, defined as L^s_{ijt}, for country j, industry i in year t. Skills, s come in three varieties, High,

[5]This sample period was selected so that it includes China, which is first observable in the WIOD dataset starting in the year 2000. Figure 5 plots averages that are weighted by employment

Medium and Low $s \in (H, M, L)$. The main independent variable of interest is the upstream global value chain length, V_{ijt}. The global value chain length captures the weighted average value chain length of trade with all foreign partners. If value chain lengths of ones trading partners reduce transaction costs and facilitate trade then we would expect the relationships between global value chain length and domestic skills to have differential effects over levels of economic development. Economic development is measured by the World Bank and is defined here as D_j. This is a high-income development dummy that is equal to one if country j is a high-income country as defined by the World Bank Development Indicators database. Industry heterogeneity is α_i, and country heterogeneity is omitted for now because it is collinear with the economic development indicator D_j. Country heterogeneity is fully accounted for within a fixed effects model in the Appendix. We also show robustness within the OLS model, by adding a country fixed effect to the model at the cost of dropping the development dummy. The matrix $\boldsymbol{X_{ijt}}$ is a set of other controls that include the real capital stocks per hour of labor employment, the skill premium, labor productivity, global value chain participation and a Herfendahl index for intermediate good concentration. All of these are interacted with the economic development indicator in our baseline model. The Herfendahl index for intermediate goods concentration is included to control for the complexity of intermediates that are one step upstream. This is because skills may be more concentrated depending on the immediate upstream sourcing complexity. But we are interested in the industrial characteristics of linkages that are farther upstream from the producer than the first stage. Thus the Herfendahl index adds such control to the regression model.

Equation 9 is estimated individually for each of the three skill categories. Table 5 reports High skilled labor, Table 6 reports medium skill and Table 7 reports results for the low skill share. This same procedure is replicated using a fixed effects estimator. In the fixed effects model, the development indicator is dropped but the interaction terms are kept and are the key coefficients of interest. Table 18, in the Appendix, reports fixed effects estimators for High skilled labor, Table 19 reports medium skill and Table 20 reports results for the low skill share.

We find that longer upstream global value chains are associated with high skilled labor in high income countries more than in medium income countries. In medium income countries, industries with longer value chains are more skill intensive in medium skilled labor. In all levels of development, longer value chains are negatively associated with low skilled workers. The summary of all of these results is our key point; value chains are correlated in upskilling both medium and high income countries but the patterns of upskilling are stronger in high income countries. The next subsection has similar conclusions using and Instrumental Variables approach. These findings are consistent with longer value chains, themselves having a higher skill requirement that is dampened in lower income countries. A dampening effect can happen if value chains facilitate trade that is itself biased by relative

Table 5: High Skill Labor Share - OLS

	(1)	(2)	(3)	(4)	(5)	(6)
dep. var: High Skill Share						
Chain Length	-0.0067	0.0117***	-0.0151***	0.0163***	-0.0030	0.0235***
	(0.0047)	(0.0040)	(0.0050)	(0.0040)	(0.0051)	(0.0038)
$\text{Country}_{High} \times$ Length	0.0327***		0.0614***		0.0546***	
	(0.0062)		(0.0067)		(0.0065)	
Herfendahl Index						
	(0.0271)					
(high income) $\times$ Herf.	0.0406					
	(0.0320)					
Participation	-0.0715***	-0.0552***	-0.1023***	-0.0489***		
	(0.0085)	(0.0067)	(0.0090)	(0.0068)		
(high income) $\times$ Participation	0.0361***		0.0655***			
	(0.0089)		(0.0097)			
Capital	0.0106***	-0.0117***	0.0117***	-0.0136***	0.0098***	-0.0144***
	(0.0015)	(0.0013)	(0.0015)	(0.0014)	(0.0015)	(0.0014)
(high income) $\times$ Capital	-0.0391***		-0.0427***		-0.0421***	
	(0.0026)		(0.0027)		(0.0027)	
Productivity	-0.0041**	0.0224***	-0.0072***	0.0255***	-0.0036**	0.0268***
	(0.0018)	(0.0017)	(0.0017)	(0.0016)	(0.0017)	(0.0016)
(high income) $\times$ Productivity	0.0451***		0.0633***		0.0611***	
	(0.0031)		(0.0029)		(0.0029)	
Skill Premium	0.0145***	-0.0144***				
	(0.0016)	(0.0018)				
(high income) $\times$ Skill Premium	-0.0676***					
	(0.0023)					
(high income)	0.0956***	0.0500***	-0.2145***	0.0681***	-0.1738***	0.0640***
	(0.0277)	(0.0029)	(0.0276)	(0.0023)	(0.0264)	(0.0024)
Constant	-0.0050	0.0401**	0.1059***	-0.0288*	0.0463**	-0.0564***
	(0.0206)	(0.0182)	(0.0210)	(0.0168)	(0.0209)	(0.0162)
Observations	6451	6451	6454	6454	6454	6454
R^2	0.449	0.296	0.332	0.280	0.322	0.274
Adjusted R^2	0.446	0.293	0.328	0.277	0.319	0.270

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

comparative advantage from factor endowments. In other words, if longer value chains enhance trade flows but do not change production technologies or endowments, then we should expect production to increase according to relative endowments. For example, an increase in lower skill production flowing to countries with an abundance of lower skilled workers.

Table 6: Medium Skill Labor Share - OLS

	(1)	(2)	(3)	(4)	(5)	(6)
dep. var: Medium Skill Share						
Chain Length	0.1650***	0.1069***	0.1162***	0.0963***	0.1159***	0.0690***
	(0.0137)	(0.0079)	(0.0129)	(0.0080)	(0.0116)	(0.0074)
(high income) × Length	-0.0438***		-0.0455***		-0.0838***	
	(0.0153)		(0.0145)		(0.0132)	
Herfendahl Index	-0.3179***					
	(0.0550)					
(high income) × Herf.	-0.1494**					
	(0.0624)					
Participation	0.0116	0.2009***	-0.0096	0.1854***		
	(0.0427)	(0.0174)	(0.0417)	(0.0174)		
(high income) × Participation	0.2311***		0.2576***			
	(0.0443)		(0.0431)			
Capital	-0.0110**	-0.0260***	-0.0056	-0.0212***	-0.0083*	-0.0184***
	(0.0047)	(0.0040)	(0.0045)	(0.0038)	(0.0043)	(0.0038)
(high income) × Capital	-0.0451***		-0.0495***		-0.0386***	
	(0.0080)		(0.0075)		(0.0073)	
Productivity	0.0186***	0.0374***	0.0152***	0.0297***	0.0183***	0.0247***
	(0.0061)	(0.0046)	(0.0058)	(0.0045)	(0.0054)	(0.0044)
(high income) × Productivity	0.0112		0.0014		-0.0042	
	(0.0085)		(0.0083)		(0.0081)	
Skill Premium	0.0248***	0.0336***				
	(0.0038)	(0.0034)				
(high income) × Skill Premium	0.0100					
	(0.0072)					
(high income)	0.4471***	0.2150***	0.4682***	0.1732***	0.6693***	0.1888***
	(0.0747)	(0.0098)	(0.0672)	(0.0080)	(0.0592)	(0.0077)
Constant	-0.3099***	-0.2072***	-0.0912	-0.0466	-0.1401***	0.0580*
	(0.0647)	(0.0377)	(0.0570)	(0.0352)	(0.0487)	(0.0332)
Observations	6451	6451	6454	6454	6454	6454
R^2	0.200	0.174	0.182	0.161	0.160	0.148
Adjusted R^2	0.195	0.170	0.177	0.157	0.155	0.144

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.1 The Effect of Value Chain Lengthening on Skills

The previous section showed that higher skills are correlated with longer global value chains. But the direction of causality is still unclear since causality can run in both directions. This section uses an instrumental variables approach to make quasi-causal inference. The strategy is as as follows. Imports into China are influenced by Chinese import tariffs. Falling Chinese import tariffs have the effect of increasing foreign access to intermediate goods that are used in Chinese manufacturing. Kee

Table 7: Low Skill Labor Share - OLS

	(1)	(2)	(3)	(4)	(5)	(6)
dep. var: Low Skill Share						
Chain Length	-0.1595***	-0.1193***	-0.1013***	-0.1130***	-0.1133***	-0.0930***
	(0.0156)	(0.0086)	(0.0152)	(0.0088)	(0.0142)	(0.0082)
(high income) × Length	0.0120		-0.0160		0.0291*	
	(0.0174)		(0.0170)		(0.0159)	
Herfendahl Index	0.2131***					
	(0.0672)					
(high income) × Herf.	0.1066					
	(0.0759)					
Participation	0.0597	-0.1453***	0.1133**	-0.1359***		
	(0.0458)	(0.0187)	(0.0449)	(0.0187)		
(high income) × Participation	-0.2669***		-0.3244***			
	(0.0471)		(0.0462)			
Capital	0.0004	0.0377***	-0.0062	0.0348***	-0.0016	0.0326***
	(0.0053)	(0.0043)	(0.0050)	(0.0042)	(0.0048)	(0.0042)
(high income) × Capital	0.0845***		0.0926***		0.0810***	
	(0.0087)		(0.0082)		(0.0080)	
Productivity	-0.0149**	-0.0602***	-0.0083	-0.0554***	-0.0150**	-0.0518***
	(0.0068)	(0.0049)	(0.0065)	(0.0048)	(0.0060)	(0.0047)
(high income) × Productivity	-0.0559***		-0.0644***		-0.0565***	
	(0.0095)		(0.0091)		(0.0088)	
Skill Premium	-0.0403***	-0.0199***				
	(0.0042)	(0.0037)				
(high income) × Skill Premium	0.0584***					
	(0.0077)					
(high income)	-0.5504***	-0.2654***	-0.2545***	-0.2408***	-0.4971***	-0.2522***
	(0.0836)	(0.0108)	(0.0765)	(0.0087)	(0.0694)	(0.0085)
Constant	1.3238***	1.1728***	0.9869***	1.0780***	1.0962***	1.0013***
	(0.0728)	(0.0408)	(0.0658)	(0.0384)	(0.0587)	(0.0366)
Observations	6451	6451	6454	6454	6454	6454
R^2	0.288	0.260	0.277	0.256	0.262	0.250
Adjusted R^2	0.284	0.256	0.273	0.253	0.258	0.247

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

and Tang (2015) show that Chinese manufacturing subsequently substituted out of foreign goods and into domestic varieties. Thus the import tariff reductions had a direct impact on both the Chinese domestic and Chinese global value chains. Production chains in China became longer during this industrial transformation. This lengthening of Chinese value chains increase the global value chain length of trade partners who import intermediates from China. Thus a decreasing import tariff in China is used as an exogenous shock to global value chain length of Chinese trading partners who are downstream of Chinese production.

Chinese import tariff reductions are constructed at the industry level by using product level TRAINS data. Product specific Chinese import tariffs, denoted as τ_{gijt} are available from the TRAINS database at the HS-6 digit level[6]. We create a country-industry import weighted average tariff for each of China's upstream trading partners.

These tariff reductions transmit heterogeneously to China's downstream partners based on both country and industry characteristics. At the country level, we consider the geographic distance from China as influencing country level sensitivity to Chinese markets. Countries that are geographically closer to China are more heavily influenced by Chinese market fluctuations. At the industry level, we consider goods that are more tradable as being more sensitive to this Chinese shock than goods that are less easily traded. Thus our instrument is a triple interaction between Chinese import tariffs, country distance from China and tradability of an industry.

To formalize ideas, defined b_j, as the great circle distance between capital's of country j and Beijing[7]. This measure is then normalized by dividing by the maximum distance from Beijing. Define the normalized measure as d_j, which is bounded by 0 and 1.

Tradability is an exogenous characteristic of the goods within the industry. Some industries such as Electrical and Optical Equipment is easier to trade than Construction or Health Care Services. Our measure of tradability uses gross exports instead of the value added measure since both measures give the same industrial rankings, Bykova and Stollinger (2017). But we do make one modification. The common measure sums across both countries and time for each industry but this time dimension can introduce endogeneity within a panel data setting. Thus we propose benchmarking an early year out of sample, in this case we benchmark the year 1995. Formally, we define the tradability for industry i, by summing exports of industry i across all countries and then dividing by the total value added in industry i;

$$T_i = \frac{\sum_j X_{i,j,t=1995}}{\sum_j VA_{i,j,t=1995}}$$

Table 8 reports the 10 most and 10 least tradable industries as well as their T_i values. We can see that the ordinal ranking of tradability follows our intuition about what we would consider tradable versus non-tradable.

The interaction of the tradability and the distance from China, defined as $S_{ij}^{-1} = (1 - T_i) \times d_j$, is a country-industry level measure of inverse sensitivity to Chinese trade. The inverse symbol in this case denotes that sensitivity decreases as the measure increases, rather than notating a mathematical

[6] We use HS0 as the base nomenclature

[7] Data is available from: http://privatewww.essex.ac.uk/ ksg/data-5.html

Table 8: Most and Least Tradable Industries

Highest Tradability			Lowest Tradability		
Rank	T_i	Industry	Rank	T_i	Industry
1	0.526	Water transport	1	0.0086	Vehicle Sale-Maint. and Retail fuel
2	0.371	Air transport	2	0.0096	Real Estate Activities
3	0.369	Elect. and Optic. Equipment	3	0.0114	Health-Social Work
4	0.299	Chemicals	4	0.0128	Construction
5	0.293	Mining and Quar.	5	0.0137	Electric. Gas and Water
6	0.277	Leather and footwear	6	0.0142	Retail trade
7	0.267	Transport Equipment	7	0.0419	Post and Telec.
8	0.257	Other Machinery	8	0.0488	Public Admin, Defense
9	0.220	Textiles	9	0.0507	Financial Intermediation
10	0.216	Other Manufacturing	10	0.0524	Hotels and Restaurants

relationship of an inverse. Both T_i and d_j are bound by zero and one. Thus we can define the sensitivity of a country industry pair by $S_{ij} = 1 - S_{ij}^{-1}$.

As a final step we define the instrumental variable, hereafter I_{ijt}, as an interaction between trade sensitivity to Chinese trade S_{ij} and Chinese tariffs on imported goods τ_{ijt}:

$$I_{ijt} = \tau_{ijt} \times S_{ij}$$

We use this instrument as an exogenous shock to global value chain length of China's trading partners using a fixed effects instrumental variable approach[8]. We run the regression independently for high, medium and low skill intensity, which is the dependent variable. It is also done independently for high income and middle income countries, thus resulting in a total of six fixed effects instrumental variable regressions. The first stage and second stage results are reported for these six regressions in Table 9. The first stage F tests show that the instrument is a significant predictor of the endogenous variable.

5 Conclusion

In this paper we traced out a link from global value chain length to demand for different kinds of labor. We hypothesized that longer value chains encourage deeper engagements in trade, which require management. In OLS regressions there is a clear association between value chain length and demand for skilled labor. We then constructed a new instrument that interacts data on Chinese import tariffs with measures of tradability and geography. The instrument varies at the country-sector level and is a good predictor of value chain length. The IV regressions increase our confidence that there is a causal relationship between value chain length and demand for different types of labor, in both high income

[8]The standard errors are robust to heteroskedasticity and autocorrelation.

Table 9: Fixed Effects IV Estimates of Chinese Value Chain Length on Skills, by Country Income

	(1) H-Income H-Skill	(2) M-Income H-Skill	(3) H-Income M-Skill	(4) M-Income M-Skill	(5) H-Income L-Skill	(6) M-Income L-Skill
GVC Length	0.389***	0.126***	0.149**	0.310**	-0.541***	-0.435***
	(0.013)	(0.012)	(0.009)	(0.031)	(0.016)	(0.039)
First Stage - GVC Length is Endogenous						
Instrument = I_{ijt}	-0.013***	-.015***	-0.013***	-.015***	-0.013***	-.015***
	0.0004	0.0013	0.0004	0.0013	0.0004	0.0013
Observations	5116	1402	5116	1402	5116	1402
F-test	945.03	121.16	945.03	121.16	945.03	121.16

Notes: H-Income denotes a high income country and M-Income denotes all countries not high income, as denoted by the World Band Development Indicator Database. Skill, defined as H-Skill="High Skill", M-Skill="Medium Skill", L-Skill="Low Skill", is measured as the share of hours worked by skill type (share in total hours), and it is the dependent variable. GVC Length is the upstream global value chain length fore each country-industry pair and the measure is taken from the Research Institute for Global Value Chains at UIBE in Beijing. Errors are robust to heteroskedasticity as well as autocorrelation. Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

and middle income countries. In the IV regressions, upstream China GVC exposure has the following effect on the skill composition of the labor force: in high income countries, strong positive effect on high skills, moderate positive effect on medium skills, and a strong negative effect on low skills. In developing countries the results are modest positive effect on both high and medium skills, and a modesty negative effect on low skills. The differential can be a driven by factor reallocation of low skilled labor from high income to middle income countries, which can dampen the upskilling effect of global value chain participation for middle income countries. The results are consistent with the idea that the expansion of GVCs has modified the usual effects of trade on the demand for factors. The GVC phenomenon by itself seems to be increasing the demand for skilled thus reducing the relative demand for unskilled labor everywhere.

A suppliers value chain length as a source of comparative advantage is a mechanism that is capable to explain these labor market effects. We show that the value chain length is a source of comparative advantage even after netting out institutional quality and controlling for product differentiation one step upstream. The rational is that longer supply chains reduce transaction costs from a blueprint maker (foreign buyer) perspective by increasing the varieties possibility frontier. More possible varieties, less need for foreign interventions, lower transaction costs, higher domestic content and skill upgrading are all possible if there is sufficient industrial depth, defined as industrial linkages or upstream value chain length.

6 Data

All measures, from other data sources, are fit to conform to WIOD industry classifications. The data is public and available at the WIOD website and readers can review Timmer, Dietzenbacher, Los, Stehrer and de Vries (2015) for technical details. We also borrow measures from Nunn (2007) as well as from Hall and Jones (1999) and are available from these authors. Value Chain Lengths are all borrowed from the Research Institute for Global Value Chains at the University of International Business and Economics in Beijing. These measures can be directly downloaded from the Institutes website. For more details, readers are directed to Wang, Wei, Yu and Zhu (2017). Bilateral tariff data comes from the UNCTAD Trade Analysis Information Systems database, which provides data at the HS-6 digit level, which were aggregated to match industries in the WIOD database. Geographical data, measuring the distance between countries came from Weidmann, Nils, Kuse, and Gleditsch (2010).

7 Appendix

7.1 Robustness of Comparative Advantage Results

Table 10: 2001

	(1)	(2)	(3)	(4)
	ri	ri	ri	ri
Supplier Chain Length	2.224***	2.015***	1.877***	1.773***
	(0.142)	(0.143)	(0.147)	(0.147)
Supplier Chain Length × inst. Quality	-1.446***	-1.364***	-0.977***	-1.015***
	(0.189)	(0.189)	(0.195)	(0.195)
(skill intensity) × (skill endowment)	11.588***	10.883***	11.176***	10.649***
	(0.369)	(0.373)	(0.372)	(0.375)
(capital intensity)×(capital endowment)	0.263***	0.326***	0.423***	0.441***
	(0.031)	(0.032)	(0.035)	(0.035)
rhhi		-1.337***		-1.186***
		(0.104)		(0.106)
Inst. × Herfindahl (benchmark)			6.461***	4.930***
			(0.675)	(0.689)
Constant	-8.366***	-7.207***	-14.379***	-11.927***
	(0.224)	(0.241)	(0.667)	(0.702)
Observations	216307	216307	216307	216307
R^2	0.420	0.420	0.420	0.420
Adjusted R^2	0.419	0.420	0.420	0.420

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: 2002

	(1)	(2)	(3)	(4)
	ri	ri	ri	ri
Supplier Chain Length	2.168***	1.899***	1.167***	1.099***
	(0.144)	(0.145)	(0.150)	(0.150)
Supplier Chain Length × inst.	-1.175***	-1.026***	0.203	0.149
	(0.192)	(0.192)	(0.201)	(0.201)
(skill intensity) × (skill endowment)	9.224***	8.388***	8.743***	8.240***
	(0.367)	(0.371)	(0.367)	(0.370)
(capital intensity) × (capital endowment)	0.606***	0.679***	0.914***	0.929***
	(0.031)	(0.031)	(0.034)	(0.034)
rhhi		-1.568***		-1.042***
		(0.102)		(0.106)
Inst. × Herfindahl (benchmark)			18.208***	16.192***
			(0.790)	(0.816)
Constant	-10.444***	-9.123***	-26.664***	-23.991***
	(0.216)	(0.233)	(0.736)	(0.784)
Observations	217700	217700	217700	217700
R^2	0.430	0.431	0.432	0.432
Adjusted R^2	0.430	0.431	0.431	0.432

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: 2003

	(1)	(2)	(3)	(4)
	ri	ri	ri	ri
Supplier Chain Length	2.708***	2.517***	1.993***	1.938***
	(0.148)	(0.148)	(0.151)	(0.151)
Supplier Chain Length × inst.	-1.889***	-1.834***	-0.932***	-0.990***
	(0.193)	(0.193)	(0.197)	(0.197)
(skill intensity) × (skill endowment)	11.127***	10.247***	9.716***	9.280***
	(0.311)	(0.317)	(0.317)	(0.320)
(capital intensity) × (capital endowment)	0.644***	0.690***	1.017***	1.010***
	(0.028)	(0.028)	(0.032)	(0.032)
rhhi		-1.611***		-1.053***
		(0.110)		(0.113)
Inst. × Herfindahl (benchmark)			13.176***	11.879***
			(0.562)	(0.579)
Constant	-10.150***	-8.607***	-22.388***	-20.176***
	(0.196)	(0.223)	(0.558)	(0.606)
Observations	217590	217590	217590	217590
R^2	0.430	0.430	0.431	0.431
Adjusted R^2	0.429	0.430	0.431	0.431

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: 2004

	(1)	(2)	(3)	(4)
	ri	ri	ri	ri
Supplier Chain Length	1.470***	1.310***	0.548***	0.499***
	(0.151)	(0.152)	(0.155)	(0.156)
Supplier Chain Length × inst. Quality	-0.388**	-0.334*	0.825***	0.803***
	(0.197)	(0.197)	(0.203)	(0.203)
(skill intensity)×(skill endowment)	9.928***	9.440***	8.292***	8.093***
	(0.280)	(0.284)	(0.288)	(0.290)
(capital intensity)×(capital endowment)	0.716***	0.748***	1.118***	1.119***
	(0.030)	(0.030)	(0.034)	(0.034)
rhhi		-1.099***		-0.603***
		(0.113)		(0.115)
Inst. × Herfindahl (benchmark)			14.120***	13.527***
			(0.585)	(0.596)
Constant	-10.662***	-9.576***	-23.648***	-22.507***
	(0.205)	(0.234)	(0.576)	(0.615)
Observations	218111	218111	218111	218111
R^2	0.430	0.430	0.431	0.431
Adjusted R^2	0.430	0.430	0.431	0.431

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: 2005

	(1)	(2)	(3)	(4)
	ri	ri	ri	ri
Supplier Chain Length	2.323***	2.259***	1.558***	1.547***
	(0.146)	(0.146)	(0.151)	(0.151)
Supplier Chain Length × inst.	-1.316***	-1.302***	-0.332*	-0.344*
	(0.192)	(0.192)	(0.199)	(0.199)
(skill intensity) × (skill endowment)	10.296***	9.974***	8.992***	8.889***
	(0.287)	(0.293)	(0.294)	(0.299)
(capital intensity) × (capital endowment)	0.845***	0.854***	1.135***	1.133***
	(0.029)	(0.029)	(0.032)	(0.032)
rhhi		-0.576***		-0.224**
		(0.109)		(0.111)
Inst. × Herfindahl (benchmark)			12.481***	12.259***
			(0.641)	(0.650)
Constant	-11.550***	-10.950***	-22.785***	-22.352***
	(0.202)	(0.232)	(0.611)	(0.648)
Observations	217749	217749	217749	217749
R^2	0.440	0.440	0.441	0.441
Adjusted R^2	0.440	0.440	0.441	0.441

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15: 2006

	(1)	(2)	(3)	(4)
	ri	ri	ri	ri
Supplier Chain Length	0.850***	0.850***	0.541***	0.571***
	(0.141)	(0.141)	(0.144)	(0.144)
Supplier Chain Length × inst.	0.186	0.124	0.595***	0.519***
	(0.188)	(0.188)	(0.192)	(0.193)
(skill intensity) × (skill endowment)	9.245***	8.866***	8.747***	8.573***
	(0.265)	(0.272)	(0.269)	(0.273)
(capital intensity) × (capital endowment)	0.914***	0.920***	1.087***	1.074***
	(0.028)	(0.028)	(0.032)	(0.032)
rhhi		-0.665***		-0.390***
		(0.103)		(0.108)
Inst. × Herfindahl (benchmark)			7.724***	6.972***
			(0.723)	(0.752)
Constant	-11.769***	-11.058***	-18.764***	-17.666***
	(0.200)	(0.229)	(0.685)	(0.749)
Observations	217682	217682	217682	217682
R^2	0.435	0.435	0.435	0.435
Adjusted R^2	0.434	0.434	0.435	0.435

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 16: 2007

	(1)	(2)	(3)	(4)
	ri	ri	ri	ri
Supplier Chain Length	0.492***	0.469***	-0.314**	-0.313**
	(0.140)	(0.140)	(0.142)	(0.142)
Supplier Chain Length × inst. Quality	0.295	0.280	1.344***	1.367***
	(0.186)	(0.186)	(0.190)	(0.190)
(skill intensity)×(skill endowment)	9.223***	9.028***	7.480***	7.564***
	(0.283)	(0.288)	(0.290)	(0.293)
(capital intensity)×(capital endowment)	0.974***	0.978***	1.270***	1.272***
	(0.027)	(0.027)	(0.029)	(0.029)
[1em] rhhi		-0.351***		0.193*
		(0.099)		(0.101)
Inst. × Herfindahl (benchmark)			19.275***	19.543***
			(0.698)	(0.712)
Constant	-11.449***	-11.060***	-27.831***	-28.272***
	(0.191)	(0.220)	(0.623)	(0.665)
Observations	221052	221052	221052	221052
R^2	0.449	0.449	0.451	0.451
Adjusted R^2	0.448	0.448	0.450	0.450

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 17: 2008

	(1)	(2)	(3)	(4)
	ri	ri	ri	ri
Supplier Chain Length	0.124	0.036	-0.665***	-0.685***
	(0.138)	(0.139)	(0.142)	(0.142)
Supplier Chain Length × Inst. Quality	0.931***	0.995***	1.888***	1.890***
	(0.183)	(0.183)	(0.187)	(0.187)
(skill intensity) × (skill endowment)	9.696***	9.284***	8.208***	8.046***
	(0.268)	(0.273)	(0.275)	(0.278)
(capital intensity) × (capital endowment)	0.929***	0.931***	1.309***	1.298***
	(0.028)	(0.028)	(0.033)	(0.033)
rhhi		-0.756***		-0.386***
		(0.093)		(0.094)
Inst. × HHI			15.356***	14.864***
			(0.662)	(0.673)
Constant	-11.406***	-10.629***	-24.967***	-24.135***
	(0.195)	(0.217)	(0.617)	(0.649)
Observations	218742	218742	218742	218742
R^2	0.446	0.446	0.447	0.447
Adjusted R^2	0.446	0.446	0.447	0.447

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7.2 Robustness of Labor Results

Table 18: High Skill Share - Fixed Effects Estimator

	(1)	(2)	(3)	(4)	(5)	(6)
High-Income × Length	0.057***		0.052***		0.051***	
	(0.005)		(0.005)		(0.005)	
Length	-0.032***	-0.004*	-0.027***	-0.003	-0.025***	-0.003
	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)
hhi	0.064***					
	(0.014)					
High-Income × hhi	-0.113***					
	(0.019)					
Participation	-0.009	-0.010**	-0.020**	-0.012***		
	(0.009)	(0.005)	(0.009)	(0.005)		
High-Income × Participation	-0.009		0.001			
	(0.010)		(0.010)			
Capital	-0.005**	0.002	-0.006**	0.002	-0.006***	0.001
	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
High-Income × Capital	0.009***		0.011***		0.011***	
	(0.003)		(0.003)		(0.003)	
L-Productivity	0.000	0.003***	0.001	0.004***	0.002	0.004***
	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
High-Income × L-Productivity	0.008***		0.008***		0.007***	
	(0.003)		(0.003)		(0.003)	
Skill Premium	0.004***	0.004***				
	(0.001)	(0.001)				
High-Income × Skill Premium	-0.001					
	(0.001)					
Observations	6451	6451	6454	6454	6454	6454
R^2	0.540	0.512	0.533	0.509	0.532	0.509
Adjusted R^2	0.494	0.465	0.487	0.462	0.486	0.461

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 19: Medium Skill Share - Fixed Effects Estimator

	(1)	(2)	(3)	(4)	(5)	(6)
GVC Length	0.033***	0.019***	0.026***	0.015***	0.023***	0.014***
	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)
Country_Rich=1 × GVC Length	-0.028***		-0.022***		-0.021***	
	(0.006)		(0.006)		(0.006)	
Herfendahl Index	-0.051***					
	(0.019)					
(high income) × Herf.	0.091***					
	(0.026)					
Participation	0.008	-0.002	0.046***	0.009		
	(0.012)	(0.006)	(0.012)	(0.006)		
(high income) × Participation	-0.009		-0.047***			
	(0.014)		(0.014)			
Capital	0.022***	0.002	0.023***	0.002	0.025***	0.002
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
(high income) × Capital	-0.030***		-0.032***		-0.035***	
	(0.004)		(0.004)		(0.004)	
Productivity	-0.023***	-0.021***	-0.028***	-0.024***	-0.030***	-0.024***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
(high income) × Productivity	0.000		0.004		0.005	
	(0.003)		(0.004)		(0.004)	
Skill Premium	-0.019***	-0.015***				
	(0.001)	(0.001)				
(high income) × Skill Premium	0.009***					
	(0.002)					
Observations	6451	6451	6454	6454	6454	6454
R^2	0.290	0.260	0.243	0.219	0.241	0.219
Adjusted R^2	0.220	0.188	0.169	0.143	0.167	0.143

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 20: Low Skill Share - Fixed Effects Estimator

	(1)	(2)	(3)	(4)	(5)	(6)
GVC Length	-0.002	-0.014***	-0.000	-0.011***	0.000	-0.011***
	(0.005)	(0.003)	(0.005)	(0.003)	(0.005)	(0.003)
(high income) × GVC Length	-0.027***		-0.027***		-0.027***	
	(0.006)		(0.006)		(0.006)	
Herf.	-0.015					
	(0.018)					
(high income) × Herf.	0.021					
	(0.025)					
Participation	0.005	0.009	-0.017	0.003		
	(0.011)	(0.006)	(0.011)	(0.006)		
(high income) × Participation	0.010		0.032**			
	(0.013)		(0.013)			
Capital	-0.018***	-0.003**	-0.018***	-0.004**	-0.019***	-0.004**
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
(high income) × Capital	0.022***		0.023***		0.025***	
	(0.004)		(0.004)		(0.003)	
Productivity	0.024***	0.018***	0.028***	0.020***	0.029***	0.020***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
(high income) × Productivity	-0.010***		-0.013***		-0.013***	
	(0.003)		(0.003)		(0.003)	
Skill Premium	0.012***	0.009***				
	(0.001)	(0.001)				
(high income) × Skill Premium	-0.008***					
	(0.002)					
Observations	6451	6451	6454	6454	6454	6454
R^2	0.647	0.641	0.638	0.633	0.637	0.633
Adjusted R^2	0.612	0.606	0.602	0.597	0.602	0.597

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

8 References

Antras and Helpman (2008), "Contractual frictions and global sourcing", In: Helpman, Martin and Verdier, The Organization of Firms in a Global Economy. Harvard University Press, Cambridge, pages 9-54

Beverelli, Koopman, KÃ¼mmritz, and Neumueller (2015), "Domestic Foundations of Global Value Chains", Available at SSRN: https://ssrn.com/abstract=2708980 .2708980

Bykova and Stollinger (2017), "Tradability Index: A Comprehensive Measure for the Tradability of Output", *wiiw Statistical Reports 6*, The Vienna Institute for International Economic Studies

Dixit, A. K., Stiglitz, J. E. (1977). Monopolistic Competition and Optimum Product Diversity.

The American Economic Review, 67(3), 297â308. http://www.jstor.org/stable/1831401

Fally and Hillberry (2018), "A Coasian Model of International Production Chains," CEPR Discussion Papers 13062, C.E.P.R. Discussion Papers.

Ferguson and Formai (2013), "Institutions-Driven Comparative Advantage, Complex Goods and Organizational Choice", *Journal of International Economics*, vol 90, 193-200

Hall and Jones (1999), "Why Do Some Countries Produce So Much More Output per Worker than Others", *Quarterly Journal of Economics*, 144, 83-116

Hidalgo Hausmann (2009), "The Building Blocks of Economic Complexity," CID Working Papers 186, Center for International Development at Harvard University

Hummels, Ishii and Yi (2001), "The nature and growth of vertical specialization in world trade", *Journal of International Economics*, vol. 54(1), pages 75-96

Jones, Charles I. 2011. "Intermediate Goods and Weak Links in the Theory of Economic Development." American Economic Journal: Macroeconomics, 3 (2): 1-28.

Kaufmann, Kraay and Mastruzzi, "Governance Matters III: Governance Indicators for 1996-2002", *World Bank Working Paper*, No. 3106

Kee and Tang (2015), "Domestic Value Added in Exports: theory and Firm Evidence from China", CESifo Working Paper 5614

Leontief, W.W. (1936) Quantitative Input and Output Relations in the Economic Systems of the United States. The Review of Economics and Statistics, 18, 105-125. https://doi.org/10.2307/1927837

Levchenko (2007), "Institutional Quality and International Trade", *Review of Economic Studies*, 74, 791-819

Nunn (2007), "Relationship-specificity, Incomplete Contracts, and the Pattern of Trade", *The Quarterly Journal of Economics*, 122(2), 569-600

Nunn and Trefler (2014), "Domestic Institutions as a Source of Comparative Advantage", *Handbook of International Economics*, chapter 5

Ranjan and Lee (2007), "Contract enforcement and International trade", *Economics and Politics*, 19(2), pages 191-218

Research Institute for Global Value Chains Database UIBE (2022), UIBE GVC Database, http://gvcdb.uibe.edu.cn

Sturgeon and Lester (2003) CITATION

Timmer, Dietzenbacher, Los, Stehrer and de Vries (2015), "An Illustrated User Guide to the World Input-Output Database: the Case of Global Automotive Production", *Review of International Economics*, 23: pages 575-605

Wang, Wei, Yu and Zhu (2017), "Characterizing Global Value Chains: Production Length and Upstreamness," *NBER Working Papers*, No. 23261

Weidmann, Nils, Kuse, and Gleditsch (2010), "The Geography of the International System: The CShapes Dataset", *International Interactions* 36 (1).

World Input Output Database (2016) doi: 10.34894/PJ2M1C.