Spatial Concentration of Buyer-Seller Matches in International Trade: The Role of Institutions and Infrastructure[1]

Fariha Kamal[2]
Center for Economic Studies
U.S. Census Bureau
4600 Silver Hill Road
Washington, DC 20233, U.S.A
fariha.kamal@census.gov
+1 (301) 763 4575

Asha Sundaram
School of Economics
University of Cape Town
Rondebosch, Cape Town 7701
South Africa
assundar.wk@gmail.com
+27 (0) 72 781 1813

Abstract

This paper explores the role of institutions and infrastructure in partner countries in shaping the patterns of spatial concentration of foreign suppliers that transact with U.S. importers. We find that the spatial concentration of suppliers within a country for an importer, as measured by a Herfindahl index, is decreasing in the quality of the origin country's contracting institutions and transport infrastructure. Additionally, we find that spatial concentration of suppliers is lower for larger U.S. importers. Our findings are consistent with the idea that there might be a greater role for networks among trading firms that operate within defined geographic boundaries, in surmounting higher costs of matching imposed by weak institutions and infrastructure.

JEL Classification: F1, F6, F14, R12
Keywords: exporter-importer match, contract enforcement, infrastructure

[1] Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. We thank Mark Kutzbach and participants of the lunch seminar at the Center for Economic Studies, U.S. Census Bureau for their invaluable comments. All errors and omissions remain our own.

[2] Corresponding author.

1. Introduction

It has long been established that firms gain productivity advantages by locating in high-density geographic areas to exploit Marshallian externalities of labor pooling, input sharing, and knowledge spillovers (Marshall, 1890). Knowledge spillovers can be especially important for exporting firms who may learn about foreign destination-specific (Koenig, 2009; Koenig, Mayneris, and Poncet, 2010), as well as buyer-specific information from neighboring exporters (Kamal and Sundaram, 2014). Using data on exporters within a particular country, or on exporter-importer pairs within a single bilateral trade relationship, these studies reveal that firms whose neighbors export are more likely to export to the same destination country, or to the same importer, respectively. While this suggests clustering of exporters selling to particular destinations, or even to specific importers, little is known in the literature about the patterns of spatial distribution of exporters across countries for individual importers.

In this study, we analyze the spatial concentration of exporters supplying to U.S. importers across all partner countries that trade with the U.S. We first present several stylized facts on the nature of supplier concentration across countries and its evolution over time. Then, we analyze the relationship between the quality of institutions and infrastructure in partner countries, key U.S. importer characteristics (size and age), and the pattern of spatial concentration of suppliers.

We argue that this question is potentially important for both importers and exporters. Importers benefit from access to suppliers who are productive, and can hence supply at lower cost. Weak institutions and infrastructure in a country that restrict access to suppliers in particular regions, or that increase the cost of matching and maintaining a trade relationship with suppliers, can erode these benefits. Similarly, weak institutions and infrastructure can affect the competitiveness of suppliers, making them less attractive to potential foreign buyers. This is of

particular concern to developing countries, where national governments are concerned with reducing regional inequality, and regional governments grapple to generate jobs and spur growth, and view exporting as a means to achieve these objectives.

We use confidential U.S. customs data maintained by the U.S. Census Bureau on merchandise import transactions between 2004 and 2011. We observe individual U.S. importer's imports from each supplier across countries, and the city in which each supplier is located. We construct a normalized Herfindahl index of concentration to measure spatial concentration of suppliers. For each U.S. importer and country it sources from, we calculate the share of suppliers in a city in the total number of suppliers. A high value for the index indicates a high concentration of suppliers in one geographic region (a city) within a country.

We find that there is substantial variation in U.S. supplier concentration across high-, middle- and low-income origin countries. Regression analyses that explores potential sources of such variation reveals that supplier concentration is higher in developing economies, where physical infrastructure is weaker, economic activity is concentrated in a small area, information on suppliers may be harder to obtain, and where contract enforcement is less perfect. Additionally, we find that larger U.S. importers tend to be more spatially diversified in their sourcing strategies. Our results are robust to controlling for unobserved year shocks, time-invariant country and importer characteristics. We also ensure that our results are not qualitatively sensitive to different measures of key variables or idiosyncrasies in the data.

Collectively, our results suggest that U.S. importers source from suppliers in close spatial proximity to each other in countries with weaker institutions and infrastructure. Suppliers to U.S. buyers are less spatially concentrated in countries with superior transport facilities. This finding resonates with the literature that examines how better transport infrastructure can facilitate

exports from remote regions (Coşar and Demir, 2014; Carballo, Cusolito, and Volpe Martincus, 2012; Duranton, 2014). Our results are also consistent with the idea that supplier concentration is higher in origin countries where institutions related to contract enforcement and access to information are weak. In such countries, it is likely that information on suppliers is harder to come by, and a potential hold-up of supplies or a delay in delivery is more likely to occur, and also costly to fix. We posit that under these circumstances, importers and exporters might rely on institutions like networks of existing suppliers operating within geographic boundaries, to learn about potential partners, their requirements, capabilities, and reliability. Networks can hence lower the cost of matching and maintaining a trade relationship, which can be higher in countries with weaker institutional environments. Kamal and Sundaram (2014) explore this idea in greater detail, focusing on U.S. buyers and Bangladeshi sellers of textile products, and find that the presence of a network of suppliers already transacting with a particular U.S. importer increases the likelihood of a Bangladeshi firm exporting to that same importer.

Finally, our work is related to the literature on institutional quality and trade patterns (Antràs and Helpman, 2008; Nunn, 2007; Levchenko, 2007). This literature shows that the quality of contracting institutions can affect the type of goods that firms import from source countries, and the organizational structure that the firm chooses to engage in trade. Firms import goods using differentiated intermediate inputs from countries where contract enforcement is stronger, and prefer to vertically integrate, rather than outsource intermediate input production, in countries where contract enforcement is weaker. In this paper, we argue that institutions might also affect the spatial dispersion of suppliers that firms source from.

The rest of the paper is organized as follows. Section 2 presents our empirical model and identification strategy. Section 3 describes the data and measurement of key variables. Section 4 discusses the empirical findings and the final section concludes.

2. Empirical Framework

In order to analyze supplier concentration we first construct a Herfindahl index as follows.

$$h_{ijt} = \sum_{n=1}^{R_{jt}} \left(\frac{x_{ijnt}}{X_{ijt}}\right)^2 \qquad (1)$$

Here, x_{ijnt} is the number of suppliers in city n located in country j exporting to importer i at time t, X_{ijt} is the total number of suppliers exporting to U.S. importer i from country j at time t and R_{jt} is the total number of cities that U.S. importers import from in country j at time t. Hence, for each importer and country with which it trades at a given point in time, the index h_{ijt} is the sum of the square of each city's share in the number of exporters.[3] Since h_{ijt} ranges between $\frac{1}{R_{jt}}$ and 1, we then obtain a normalized index as follows, to account for the fact that number of cities varies considerably by country.

$$H_{ijt} = \left(\frac{h_{ijt} - \frac{1}{R_{jt}}}{1 - \frac{1}{R_{jt}}}\right) * 100 \qquad (2)$$

H_{ijt} may take on values between 0 and 100 including 0 and 100. A high value of H_{ijt} indicates that at time t, importer i's suppliers in country j are spatially concentrated.

[3] In the current formulation, we treat each city, within a country, as an island with no spatial relationship to other cities. We leave this extension for future work.

We seek to understand how supplier concentration varies by country and importer characteristics. To do this, we estimate the following equation.[4]

$$H_{ijt} = \alpha + \beta I_{jt} + \gamma Y_{it} + \delta_j + \mu_t + \vartheta_i + \epsilon_{ijt} \qquad (3)$$

I_{jt} refers to a set of characteristics of country j at time t and Y_{it}, a set of importer characteristics of importer i at time t. δ_j , μ_t and ϑ_i refer to a set of country, year and importer fixed-effects respectively. ϵ_{ijt} is an idiosyncratic error. I_{jt} includes measures of institutions and infrastructure at the country level and additional control variables. We are interested in a particular dimension of institutions – contract enforcement – that are pertinent for conducting business. We use the number of days required to enforce a commercial contract. Our measures of infrastructure include the number of fixed internet subscribers per 100 to indicate presence of internet technology that enables information transmission, container traffic at ports, and percentage of paved roads. We also include the percentage of the urban population in the country's largest city as a control for the density of economic activity, log of GDP per capita and the log of total population.[5] In Appendix Table A1, we present pairwise correlations between country-level independent variables to ensure that no two measures of interest (institution and infrastructure) are highly correlated.[6] Y_{it} includes an importer's size measured as total employment as well as the total number of establishments, and age.

[4] We have a simple location choice model in mind to motivate equation (3). An importer has a distribution, H_0, of suppliers in country j at time t. In some subsequent year, it may need to replace old suppliers (with some probability) and add new suppliers from the set of suppliers across all cities within a country. The choice of suppliers will be informed by existing suppliers in a city, country level characteristics such as institutions and infrastructure, supplier and buyer characteristics, and interactions between these factors. Since we analyze spatial concentration conditional on importing from a particular country, we abstract from the decision to choose the origin country. We argue that conditional on importing from a particular country, spatial concentration is negatively correlated with institutional and infrastructure quality.

[5] An alternate measure of concentration of economic activity that we use is the percentage of the population in urban agglomerations greater than one million. We find that our key results hold qualitatively.

[6] In Table A1, we find that GDP per capita is highly correlated with fixed-line internet users per 100. Hence, we test for robustness of results using alternate measures of internet usage, number of secure servers per 1 million people

We estimate several variants of equation (3). First, we exclude country, year and firm fixed-effects. Hence, we exploit all the cross-section variation across countries and firms in the institutional and infrastructure-related variables. Next, we add state, industry and year fixed-effects. Year effects account for trends in supplier concentration. It is possible that state-level policies in the U.S. facilitate importer-exporter matches in particular countries and cities. We account for such unobserved, time-invariant, state-specific factors with a fixed-effect for the state in which the importer is located. Industry fixed-effects account for industry-specific time-invariant shocks that are correlated with both concentration and independent variables of interest jointly. For instance, it is possible that some industries, like automobile manufacture, utilize technology that lends itself to spatial agglomeration, and is also more dependent on transport infrastructure. Finally, our preferred specification is the most rigorous, where we include year, country and U.S. importer fixed-effects to account for unobserved shocks correlated with institutional and infrastructure-related variables and supplier concentration jointly. For instance, it is possible that certain types of U.S. importers are more likely to diversify spatially, and are also more likely to source from advanced countries, where institutions are stronger, allowing greater diversification. Our preferred specification accounts for such factors, as long as they are time-invariant. Thus, identification of the association between institutional and infrastructure-related variables and supplier concentration is derived solely through deviations in these variables from the country mean, after having accounted for unobserved, U.S. importer-specific, time-invariant shocks.

The empirical specification above is linear although our dependent variable is a percentage and so bounded between 0 and 1. We choose a linear model in order to implement the

---

and number of internet users per 100 people, with correlation coefficients of 0.50 and 0.60, respectively. Results are discussed in Section 4.3.

high dimensional fixed effects and allow for a more straightforward interpretation of the implied marginal effects. However, as a robustness check, we employ a fractional logit model (Papke and Wooldridge, 1996) and our results remain qualitatively unchanged.[7]

Our key hypotheses are as follows. If contract enforcement is a difficult and lengthy process, learning about a supplier's capabilities and reliability via a network can be more important, in order to avoid future hold-up problems or delays in shipments. This implies that the number of days required to enforce a contract in a country is likely to be positively correlated with supplier concentration. We expect information on suppliers to be less accessible (and/or accessible at a higher cost) to U.S. importers in countries with weaker internet presence. If this is true, we hypothesize that U.S. importers would rely more on networks to identify potential suppliers. Similarly, potential suppliers might utilize networks to approach potential buyers when information flows are unreliable. Since network effects operate within geographic boundaries, lower internet presence would be correlated with greater supplier concentration.

Transport costs are much lower in countries with efficient ports and better land transport infrastructure, affording greater scope to U.S. importers for spatial diversification of suppliers. Hence, we expect the percentage of paved roads and container traffic in ports to be negatively associated with supplier concentration. In countries where economic activity is more concentrated, we anticipate lower possibilities for diversification, implying a positive relationship between the percentage of the urban population in the largest city and supplier concentration. Finally, we include the log of population to control for country size, and the log of GDP per capita to account for the role of any unobserved institutional factors that might affect our variables of interest and supplier concentration simultaneously.

[7] Results available upon request.

While characteristics of the supplier's origin country can affect supplier concentration, the level of detail in our data allows us to additionally investigate the relationship between U.S. importer characteristics and supplier concentration. We hence ask if, controlling for origin country characteristics, older and/or larger firms are more spatially diversified in their acquisition of suppliers. It is possible that larger and/or more established U.S. importers are better able to surmount barriers imposed by weak institutions and infrastructure, and have to rely less on networks to acquire suppliers.

3. Data

3.1 U.S. Census Data

We test our hypotheses using confidential U.S. firm-merchandise import transaction linked data, the Linked/Longitudinal Foreign Trade Transactions Database (LFTTD-IMP), in conjunction with country level measures obtained from the World Bank between 2004 and 2011. We restrict attention to arm's length transactions only thus abstracting from multinationals' subsidiary location choices. Every U.S. merchandise import transaction contains a unique identifier for the foreign exporter. The foreign exporter is uniquely identified by the "Manufacturer ID" (MID), a required field on Form 7501, the form U.S. importers are required to file with the U.S. Customs and Border Protection (CBP).[8]

The MID identifies the manufacturer or shipper of the merchandise by an alphanumeric code that is constructed using a pre-specified algorithm with a maximum length of 15 characters.[9] The last three characters in the MID designate the city where the manufacturer is located. Thus, we consider each distinct three-letter code as a unique city within a country. We

[8] See form http://forms.cbp.gov/pdf/cbp_form_7501.pdf.
[9] See Block 13 (pg. 7) for description of MID and Appendix 2 (pg. 30) for instructions on constructing MID at http://forms.cbp.gov/pdf/7501_instructions.pdf.

carry out robustness checks to address concerns of instances where there may be multiple cities within a country that begin with the same first three letters. We also exclude observations for countries in the sample that are associated with only one city and represent a tiny share (less than 0.01%) of the overall sample.

Information on an importer's basic firm characteristics are from the Longitudinal Business Database (LBD). The LBD consists of data on all private, non-farm U.S. establishments in existence that have at least one paid employee (Jarmin and Miranda, 2002). We use the LBD to obtain information on firm employment, age, industry, and state. For firms with multiple plants, age is calculated as the difference between the year of interest and the year of establishment of its oldest plant. The firm is considered to be operating in the industry and state where the largest share of its employment is housed.

3.2 World Bank Data

The country level measures are sourced from the World Bank's public-use databases. The measure of institution, contract enforcement, is obtained from the World Bank's Doing Business database that begins in 2004.[10] The Doing Business database contains three measures of contract enforcement. Our preferred measure, the number of days required to enforce a commercial contract, displays the most temporal variation. However, in robustness checks (Table 4) we also employ the two other measures, percentage cost of claims and number of procedures required to enforce contracts, to find that our results remain qualitatively unchanged.

We consider three main measures of infrastructure within a country from the World Development Indicators (WDI) database.[11] First, we measure internet technology presence in a country as the number of fixed broadband internet subscribers with a digital subscriber line,

[10] See http://www.doingbusiness.org/ for details on methodology.
[11] See http://data.worldbank.org/data-catalog/world-development-indicators for details on data coverage.

cable modem, or other high-speed technology per 100 people. This measure captures the ease of information transmission between buyers and sellers for instance, via company websites, portals and electronic communication in general. We also consider two alternative measures of internet technology presence in a country (Table 5), number of secure servers using encryption technology in internet transactions per 1 million people and number of internet users per 100 people, to still find broad support for our hypotheses. Second, we include a measure of port container traffic that enters our specification in logs to capture the efficiency of ports. It measures the flow of containers from land to sea transport modes in twenty-foot equivalent units. Finally, we include the percentage of roads that are paved as a share of all the country's roads, measured in length.

The remaining country level measures are also sourced from the WDI. We include the percentage share of a country's urban population living in that country's largest metropolitan area as a measure of overall economic density in the country. Lastly, we include the GDP per capita (in 2005 USD) and total population, both of which enter our specification in logs.

4. Results

4.1 Summary Statistics

In Figure 1, we present the distribution of supplier concentration for low-, middle- and high-income countries in the year 2007. In each year, we divide countries into three quantiles based on GDP per capita. See Table A2 for a list of countries in each group in 2011. The measures of institution and infrastructure are meaningfully correlated with these rankings. Higher the ranking, lower (higher) the cost of contract enforcement (quality of infrastructure), thus providing an intuitive way to preview our findings. We find that the distribution of the

supplier concentration measure in high-income countries lies farthest to the left, followed by the distribution for middle-income countries, and then low-income countries to the right. This indicates that overall, U.S. supplier concentration in high-income and middle-income countries is lower than in low-income countries. This hierarchy is consistent with the hypothesis that stronger institutions and infrastructure can support spatial diversification of supplier matches.

In Table 1, we look at average supplier concentration indices for high-, middle- and low-income countries for three years, 2004, 2007 and 2011. We find that in all three years, mean concentration indices for high- and middle-income countries, and for high- and low-income countries are significantly different from each other. In addition, while concentration indices for all three country groups decline over time, the negative trend is stronger for high-income economies than for middle- and low-income economies. This suggests that middle- and low-income countries do not completely catch-up to low concentration ratios in high-income countries, in spite of developments in the quality of institutions and infrastructure that one might expect in fast-growing, emerging, middle- and low-income economies over time.

Table 2, panels A and B, presents the mean and standard deviations of the country and importer level variables, respectively. In the average country in our sample, it takes about 622 days to enforce a commercial contract; 1 out of 20 people have access to fixed broadband internet technology; 53% of roads are paved while 34% of the urban population lives in the largest metropolitan city. The average importer in our sample has been in operation for about 14 years employing about a 190 workers across 5 establishments. In addition, Panel B shows that an average importer sources from about 4 suppliers per country and from 2 cities per country with a total of about 11 suppliers, in 6 cities, across 3 countries.

4.2 Baseline results

Table 3 presents our baseline regression results based on variations of specification (3). Column (1) presents results without fixed-effects. Column (2) includes year, industry and state fixed-effects. Columns (3) and (4) include U.S. importer, country and year fixed-effects. In column (4), the dependent variable is an import-value weighted measure of the Herfindahl index. Our results offer strong support for our hypotheses. Across all specifications, the number of fixed internet subscribers per 100, the percentage of paved roads and container traffic in ports are all negative and statistically significant, indicating a relationship between weak infrastructure, potentially lower access to information, and greater supplier concentration. Results suggest that a 10 percentage-point increase in the percentage of fixed internet subscribers in a country is associated with a one-half percentage point decrease in the supplier concentration index. A 10 percentage-point increase in the percentage of roads paved is associated with a 1.6 percentage point decrease in the supplier concentration index. Finally, a one percent increase in container traffic is associated with a one percent increase in the concentration index.

The number of days to enforce a commercial contract, and the percentage of the urban population in the largest city are both positive and significant, suggesting that weak contract enforcement and concentration of economic activity are associated with greater supplier concentration. We find that if the number of days required to enforce a contract increases by a 100, the concentration index increases by 1.5 percentage points. Finally, we find that supplier concentration is lower for larger U.S. importers, suggesting that large importers might have higher tolerance for the barriers imposed by weak institutions and infrastructure in sourcing from less-connected regions.

4.3 Robustness checks

Our results are robust to alternative measures of contract enforcement and internet technology availability as shown in Tables 4 and 5. Since we identify cities within a country using the three letter codes extracted from the MID, it is possible that for cities that begin with the same three letters, a single code many actually represent multiple cities and introduce measurement bias in our spatial concentration index. However, as long as the incidence of such cases are not systematically correlated with our measures of country institution and infrastructure, our coefficient estimates of these measures of interest will remain unbiased. However, as an additional robustness check, we exclude large countries from our sample both in terms of population (Table 6) and in terms of land area (Table 7), with the premise that larger countries are likely to have larger number of cities that may share the same first three letters. Our results retain their flavor when large countries are excluded.[12]

Finally, it is possible that the MID represents a shipping agent or other direct suppliers such as intermediaries. If all MIDs in the sample represented shippers instead of manufactures, spatial concentration of suppliers may be driven by the fact that manufacturers across the origin country use a limited number of large shipping agents located in a port city or a major hub. We hence estimate our main specification on a restricted sample of imports in textile products only. Due to strict rules-of-origin requirements, the MID for textile shipments represents the "manufacturer" as defined in Title 19 Code of Federal Regulations (CFR), i.e. "the entity performing the origin-conferring operations". The definition of what constitutes the foreign supplier is less strict for non-textile products and many include both manufacturers and shippers.[13] Textile products include both textile or apparel products as defined under Section

[12] An exception is the internet infrastructure variable, which loses significance and flips sign. However, using the alternate measures of internet presence yield results similar to our baseline.
[13] See http://www.gpo.gov/fdsys/pkg/CFR-2011-title19-vol1/pdf/CFR-2011-title19-vol1-sec102-23.pdf.

102.21, Title 19, CFR[14], classified as any products in two-digit HS codes 50 through 63.[15] We find that our results are qualitatively unchanged in this restricted sample, as shown in Table 8.

5. Conclusion

This study tests the role of a partner country's contracting environment and infrastructure in shaping the patterns of spatial concentration of suppliers to U.S. importers. We also examine the characteristics of U.S. importers associated with our measure of spatial concentration. We find that the spatial concentration is lower in countries with better contract enforcement and infrastructure quality, as well as for larger importers. Collectively, these findings are consistent with the idea that there is a role for networks that operate within defined geographic spaces in surmounting higher costs of matching imposed by weak institutions and infrastructure. We note here that there may be other mechanisms driving the spatial concentration of suppliers. For instance, if importers need to employ multiple agents to monitor their suppliers, then they may prefer to source from geographically proximate exporters in order to reduce monitoring costs. This might result in concentration. Additionally, contracting institutions might interact with transport infrastructure to drive spatial concentration of suppliers. Better transport infrastructure would lower monitoring costs, since travel to multiple cities would be less expensive. Disentangling these possible channels offers scope for future work.

[14] See http://www.gpo.gov/fdsys/pkg/CFR-2011-title19-vol1/pdf/CFR-2011-title19-vol1-sec102-21.pdf.
[15] See http://hts.usitc.gov/ for details on each HS chapter.

References

Antràs, Pol, and Elhanan Helpman. 2008. "Contractual Frictions and Global Sourcing." In *The Organization of Firms in a Global Economy*, eds. Elhanan Helpman, Dalia Marin, and Thierry Verdier, 9-54. Cambridge, MA: Harvard University Press.

Carballo, Jeronimo, Ana Cusolito, and Christian Volpe Martincus, "Routes, Exports, and Employment in Developing Countries: Following the Trade of the Inca Roads." Unpublished.

Coşar, Kerem A., and Banu Demir. 2014. "Domestic Road Infrastructure and International Trade: Evidence from Turkey." Unpublished.

Duranton, Gilles. 2014. "Roads and Trade in Colombia." Unpublished.

Jarmin, Ron, S., and Javier Miranda. 2002. "The Longitudinal Business Database." U.S. Census Bureau Center for Economic Studies Working Paper 02-17.

Kamal, Fariha, and Asha Sundaram. 2014. "Buyer-Seller Relationships in International Trade: Do your Neighbors Matter?" U.S. Census Bureau Center for Economic Studies Working Paper 14-42.

Koenig, Pamina. 2009. "Agglomeration and the Export Decisions of French Firms." *Journal of Urban Economics* 66: 186-195.

Koenig, Pamina, Florian Mayneris, and Sandra Poncet. 2010. "Local Export Spillovers in France." *European Economic Review* 54: 622-641.

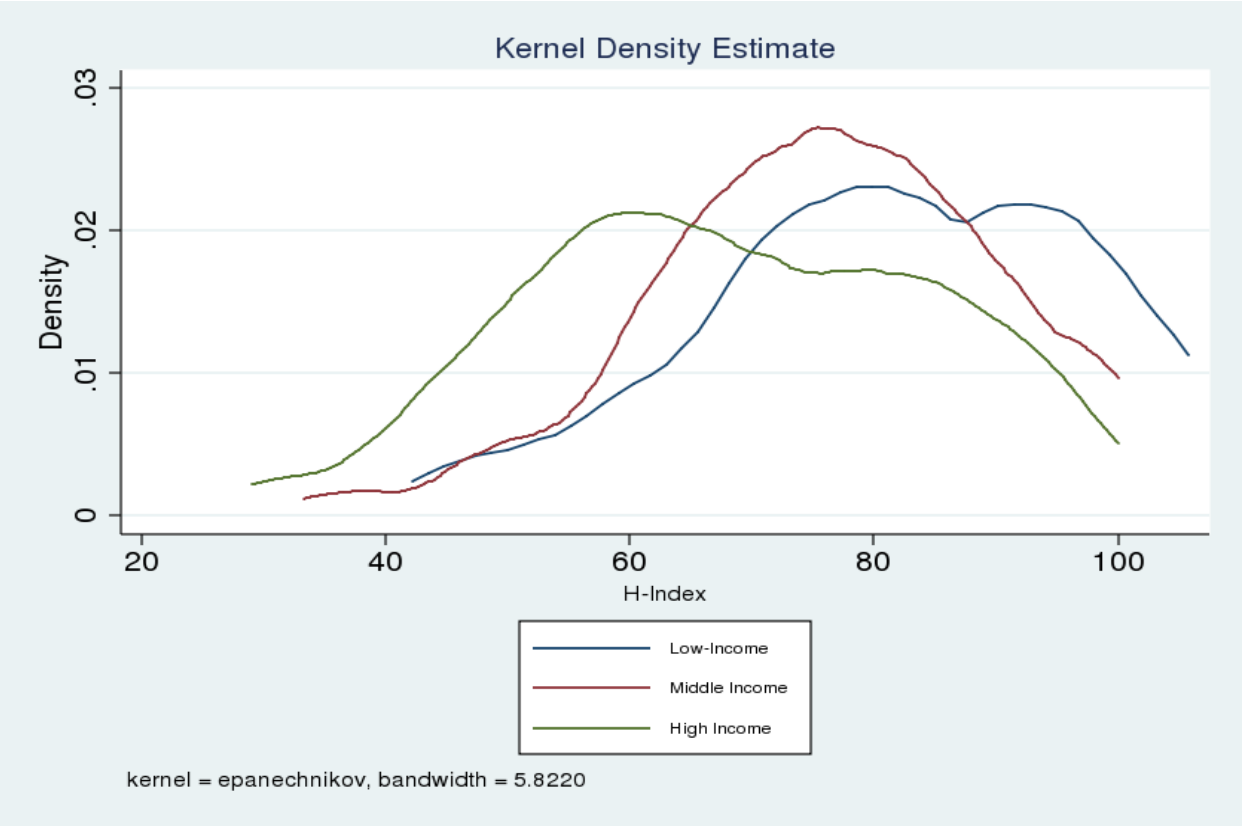
Levchenko, Andrei. 2007. "Institutional Quality and International Trade." *Review of Economic Studies* 74(3): 791-819.

Marshall, A. 1890. *Principles of Economics*. London: Macmillan.

Nunn, Nathan. 2007. "Relationship-Specificity, Incomplete Contracts and the Pattern of Trade." *Quarterly Journal of Economics* 122(2): 569-600.

Papke, Leslie, E., and Jeffrey M. Wooldridge. 1996. "Econometric Methods for Fractional Response Variables with an Application to 401(K) Plan Participation Rates." *Journal of Applied Econometrics* 11: 619-632.

Figure 1. Distribution of H_{ijt} by country groupings, 2007



Notes: H_{ijt} is a Herfindahl index of supplier concentration that varies by importer and origin country.

Table 1. Average H_{ijt}, Selected Years

	2004	**2007**	**2011**
Low Income	81.28	81.56	79.47
Middle Income	77.09	75.99	75.15
High Income	68.78	67.99	65.58
	T-statistics of difference in means		
Middle vs. Low	1.40	2.04	1.52
High vs. Low	4.24	5.02	4.91
High vs. Medium	2.82	2.95	3.37

Notes: H_{ijt} is a Herfindahl index of supplier concentration that varies by importer and origin country.

Table 2. Summary Statistics

Panel A Country Level

Variables	Mean (sd.)
Cost of contract enforcement (number of days)	621.68
	(306.32)
Fixed Internet Subscribers (per 100 people)	4.91
	(8.90)
Log Container Traffic	13.97
	(1.64)
Paved Roads (%)	52.94
	(33.21)
Population in Largest City (%)	33.87
	(17.66)
Log GDP per Capita	8.04
	(1.63)
Log Total Population	15.10
	(2.30)

Panel B Importer Level

Variables	Mean (sd.)
Age	13.52
	(10.82)
Employment	193.08
	(3,982.30)
Number of Establishments	4.67
	(92.09)
Number of foreign suppliers/country	4.35
	(74.20)
Number of foreign suppliers	11.12
	(469.12)
Number of cities/country	2.25
	(7.52)
Number of cities	5.74
	(44.86)
Number of countries	2.55
	(3.71)

Table 3. Determinants of Spatial Concentration of Foreign Suppliers to U.S. Importers, 2004-2011.

Dependent variable:	(1) H_{ijt}	(2) H_{ijt}	(3) H_{ijt}	(4) H_{ijt} *(Value Weighted)*
Country Characteristics				
Cost of Contract Enforcement (Days)	0.004***	0.003***	0.015***	0.012***
	(0.000)	(0.000)	(0.001)	(0.001)
Fixed Internet Subscribers (per 100 people)	-0.104***	-0.115***	-0.052***	-0.055***
	(0.004)	(0.005)	(0.015)	(0.013)
Log Container Traffic	-1.779***	-1.741***	-1.354***	-1.383***
	(0.030)	(0.032)	(0.282)	(0.250)
Paved Roads (%)	-0.011***	-0.006***	-0.161***	-0.149***
	(0.001)	(0.001)	(0.014)	(0.013)
Population in Largest City (%)	-0.045***	-0.046**	0.406***	0.335***
	(0.002)	(0.002)	(0.099)	(0.088)
Log GDP per Capita	0.475	-0.231***	-6.319***	-5.080***
	(0.048)	(0.054)	(0.624)	(0.553)
Log Total Population	-3.839***	-4.277***	0.712	-0.315
	(0.042)	(0.045)	(2.224)	(1.973)
Importer Characteristics				
Log Employment	-1.855***	-2.848***	-2.256***	-1.915***
	(0.018)	(0.019)	(0.062)	(0.055)
Log Number of Establishments	-2.010***	-1.320***	-1.057***	-0.873***
	(0.027)	(0.031)	(0.150)	(0.131)
Log Firm Age	0.153***	0.525***	0.611***	0.569***
	(0.027)	(0.028)	(0.114)	(0.101)
Industry FE	-	Y	-	-
State FE	-	Y	-	-
Year FE	-	Y	Y	Y
Country FE	-	-	Y	Y
Firm FE	-	-	Y	Y
Observations	1,427,898	1,427,458	1,427,898	1,434,132
Adjusted R-squared	0.10	0.13	0.33	0.32

Notes: Robust standard errors in parentheses. Significance level: * if $p < 0.10$, ** if $p < 0.05$; *** if $p < 0.01$. H_{ijt} is a Herfindahl index of supplier concentration that varies by importer and origin country.

Table 4. Alternative Measures of Contract Enforcement.

Dependent variable: H_{ijt}	(1) % of Cost	(2) Number of Procedures
Country Characteristics		
Cost of Contract Enforcement	0.040†	0.297**
	(0.025)	(0.0997)
Fixed Internet Subscribers (per 100 people)	-0.080***	-0.068***
	(0.016)	(0.015)
Log Container Traffic	-2.117***	-2.162***
	(0.275)	(0.271)
Paved Roads (%)	-0.139***	-0.138***
	(0.014)	(0.014)
Population in Largest City (%)	0.310***	0.343**
	(0.101)	(0.099)
Log GDP per Capita	-6.678***	-6.901***
	(0.624)	(0.630)
Log Total Population	1.092	2.366
	(2.352)	(2.219)
Importer Characteristics		
Log Employment	-2.261***	-2.261***
	(0.062)	(0.062)
Log Number of Establishments	-1.065***	-1.066***
	(0.150)	(0.150)
Log Firm Age	0.638***	0.636***
	(0.114)	(0.114)
Year FE	Y	Y
Country FE	Y	Y
Firm FE	Y	Y
Observations	1,427,898	1,427,898
Adjusted R-squared	0.33	0.33

Notes: Robust standard errors in parentheses. Significance level: if † $p < 0.15$, * if $p < 0.10$, ** if $p < 0.05$; *** if $p < 0.01$. H_{ijt} is a Herfindahl index of supplier concentration that varies by importer and origin country.

Table 5. Alternative Measures of Internet Technology Presence.

Dependent variable: H_{ijt}	(1) Secure Servers per 1 million people	(2) Internet Users per 100 people
Country Characteristics		
Cost of Contract Enforcement (Days)	0.014***	0.015***
	(0.001)	(0.001)
Information Infrastructure	-0.001***	-0.011†
	(0.000)	(0.007)
Log Container Traffic	-1.495***	-1.122***
	(0.279)	(0.273)
Paved Roads (%)	-0.183***	-0.174***
	(0.013)	(0.014)
Population in Largest City (%)	0.310***	0.398***
	(0.101)	(0.098)
Log GDP per Capita	-5.997***	-5.275***
	(0.552)	(0.539)
Log Total Population	-0.679	2.238
	(2.197)	(2.154)
Importer Characteristics		
Log Employment	-2.253***	-2.265***
	(0.062)	(0.062)
Log Number of Establishments	-1.046***	-1.082***
	(0.150)	(0.150)
Log Firm Age	0.596***	0.620***
	(0.113)	(0.113)
Year FE	Y	Y
Country FE	Y	Y
Firm FE	Y	Y
Observations	1,423,631	1,431,764
Adjusted R-squared	0.33	0.33

Notes: Robust standard errors in parentheses. Significance level: if † $p < 0.15$, * if $p < 0.10$, ** if $p < 0.05$; *** if $p < 0.01$. H_{ijt} is a Herfindahl index of supplier concentration that varies by importer and origin country.

Table 6. Exclude large countries, by population size.

Dependent variable: H_{ijt}	(1) Top 2	(2) Top 3	(3) Top 5
Country Characteristics			
Cost of Contract Enforcement (Days)	0.015***	0.015***	0.016***
	(0.001)	(0.001)	(0.001)
Fixed Internet Subscribers (per 100 people)	0.015	0.015	0.005
	(0.017)	(0.017)	(0.018)
Log Container Traffic	-1.566***	-1.652***	-1.682***
	(0.291)	(0.290)	(0.299)
Paved Roads (%)	-0.117***	-0.122***	-0.117***
	(0.024)	(0.024)	(0.026)
Population in Largest City (%)	0.376**	0.286**	0.322***
	(0.099)	(0.099)	(0.102)
Log GDP per Capita	6.835***	5.772***	6.348***
	(1.023)	(1.033)	(1.081)
Log Total Population	-7.622***	-6.326**	-4.833**
	(2.222)	(2.213)	(2.243)
Importer Characteristics			
Log Employment	-1.846***	-1.794***	-1.843***
	(0.076)	(0.077)	(0.080)
Log Number of Establishments	-1.159***	-1.129***	-1.179***
	(0.160)	(0.162)	(0.160)
Log Firm Age	-0.199	-0.197	-0.224
	(0.141)	(0.144)	(0.148)
Year FE	Y	Y	Y
Country FE	Y	Y	Y
Firm FE	Y	Y	Y
Observations	960,427	923,277	874,390
Adjusted R-squared	0.26	0.26	0.26

Notes: Robust standard errors in parentheses. Significance level: * if $p < 0.10$, ** if $p < 0.05$; *** if $p < 0.01$. H_{ijt} is a Herfindahl index of supplier concentration that varies by importer and origin country.

Table 7. Exclude large countries, by land area.

Dependent variable: H_{ijt}	(1) Top 2	(2) Top 3	(3) Top 5
Country Characteristics			
Cost of Contract Enforcement (Days)	0.014***	0.014***	0.015***
	(0.001)	(0.001)	(0.001)
Fixed Internet Subscribers (per 100 people)	0.018	0.019	0.017
	(0.016)	(0.016)	(0.017)
Log Container Traffic	-1.817***	-1.797***	-1.561***
	(0.288)	(0.288)	(0.292)
Paved Roads (%)	-0.123***	-0.124***	-0.119***
	(0.024)	(0.024)	(0.024)
Population in Largest City (%)	0.239**	0.240**	0.392***
	(0.097)	(0.097)	(0.101)
Log GDP per Capita	4.305***	4.186***	6.659***
	(0.914)	(0.913)	(1.033)
Log Total Population	-8.172***	-8.282***	-8.085***
	(2.223)	(2.227)	(2.245)
Importer Characteristics			
Log Employment	-1.869***	-1.866***	-1.849***
	(0.075)	(0.075)	(0.077)
Log Number of Establishments	-1.184***	-1.135***	-1.035***
	(0.158)	(0.160)	(0.164)
Log Firm Age	-0.089	-0.001	-0.112
	(0.137)	(0.139)	(0.144)
Year FE	Y	Y	Y
Country FE	Y	Y	Y
Firm FE	Y	Y	Y
Observations	1,019,270	1,003,698	925,233
Adjusted R-squared	0.25	0.26	0.26

Notes: Robust standard errors in parentheses. Significance level: * if $p < 0.10$, ** if $p < 0.05$; *** if $p < 0.01$. H_{ijt} is a Herfindahl index of supplier concentration that varies by importer and origin country.

Table 8. Alternative sample.

Dependent variable: H_{ijt}	(1) Textile Products Only
Country Characteristics	
Cost of Contract Enforcement (Days)	0.006*
	(0.004)
Information Infrastructure	-0.143***
	(0.041)
Log Container Traffic	-5.116***
	(0.740)
Paved Roads (%)	-0.105***
	(0.033)
Population in Largest City (%)	0.537**
	(0.268)
Log GDP per Capita	-1.949
	(1.524)
Log Total Population	-16.42**
	(7.611)
Importer Characteristics	
Log Employment	-2.395***
	(0.141)
Log Number of Establishments	0.0698
	(0.359)
Log Firm Age	0.518**
	(0.262)
Year FE	Y
Country FE	Y
Firm FE	Y
Observations	257,531
Adjusted R-squared	0.36

Notes: Robust standard errors in parentheses. Significance level: * if $p < 0.10$, ** if $p < 0.05$; *** if $p < 0.01$. H_{ijt} is a Herfindahl index of supplier concentration that varies by importer and origin country.

Appendix

Table A1 Pairwise correlations between country-level variables

	Cost of Contract Enforcement	Fixed Internet Subscribers	Log Container Traffic	Paved Roads	Population in Largest City	Log GDP per Capita	Log Total Population
Cost of Contract Enforcement	1						
Fixed Internet Subscribers	-0.40	1					
Log Container Traffic	-0.25	-0.14	1				
Paved Roads	-0.30	0.68	-0.03	1			
Population in Largest City	-0.26	0.28	-0.17	0.46	1		
Log GDP per Capita	-0.31	0.91	-0.28	0.58	0.33	1	
Log Total Population	0.11	-0.49	0.71	-0.39	-0.66	-0.62	1

Notes: See Section 3.2 for variable definitions.

Table A2 Countries by Income Groupings, 2011

Low Income	Middle Income	High Income
Afghanistan	Albania	Antigua and Barbuda
Bangladesh	Algeria	Australia
Benin	Angola	Austria
Bolivia	Armenia	Bahamas
Burkina Faso	Azerbaijan	Bahrain
Burundi	Belarus	Barbados
Cambodia	Belize	Belgium
Cameroon	Bhutan	Bermuda
Central African Republic	Bosnia and Herzegovina	Brunei
Chad	Botswana	Canada
Comoros	Brazil	Chile
Congo, Democratic Rep.	Bulgaria	Croatia
Cote d'Ivoire	Cabo Verde	Cyprus
Djibouti	China	Czech Republic
Egypt	Colombia	Denmark
Eritrea	Congo, Republic of	Equatorial Guinea
Ethiopia	Costa Rica	Estonia
Gambia	Dominica	Finland
Ghana	Dominican Republic	France
Guinea	Ecuador	Germany
Guinea-Bissau	El Salvador	Greece
Guyana	Fiji	Hong Kong
Haiti	Gabon	Hungary
Honduras	Georgia	Iceland
India	Grenada	Ireland
Kenya	Guatemala	Israel
Kiribati	Indonesia	Italy
Kyrgyzstan	Iran	Japan
Laos	Iraq	Kuwait
Lesotho	Jordan	Latvia
Liberia	Kazakhstan	Lebanon
Madagascar	Kosovo	Lithuania
Malawi	Libya	Luxembourg
Mali	Macedonia	Macao
Mauritania	Malaysia	Malta
Moldova	Maldives	Mexico
Mongolia	Marshall Islands	Netherlands
Mozambique	Mauritius	New Zealand
Nepal	Micronesia	Norway
Nicaragua	Montenegro	Oman
Niger	Morocco	Palau

Nigeria	Namibia	Panama
Pakistan	Paraguay	Poland
Papua New Guinea	Peru	Portugal
Philippines	Romania	Puerto Rico
Rwanda	Saint Lucia	Qatar
Sao Tome and Principe	St. Vincent and the Grenadines	Russia
Senegal	Western Samoa	St. Kitts and Nevis
Sierra Leone	Serbia	Saudi Arabia
Solomon Islands	South Africa	Seychelles
Sudan	Sri Lanka	Singapore
Tajikistan	Suriname	Slovakia
Tanzania	Swaziland	Slovenia
Timor-Leste	Thailand	South Korea
Togo	Tonga	Spain
Uganda	Tunisia	Sweden
Uzbekistan	Turkmenistan	Switzerland
Vietnam	Ukraine	Trinidad and Tobago
West Bank	Vanuatu	Turkey
Yemen	Venezuela	United Arab Emirates
Zambia		United Kingdom
Zimbabwe		Uruguay

Notes: Countries are grouped into three quantiles of gdp per capita within a year.