The Margins of Global Sourcing: Theory and Evidence from U.S. Firms*

Pol Antràs
Harvard University

Teresa C. Fort
Tuck School at Dartmouth

Felix Tintelnot
University of Chicago
and Princeton University

March 13, 2014

(PRELIMINARY and INCOMPLETE)

Abstract

This paper studies the extensive and intensive margins of firms' global sourcing decisions. First, it presents three new facts on U.S. firms' import behavior that highlight the importance of the extensive margin in explaining cross-sectional variation in U.S. import volumes. These facts motivate the development of a quantifiable multi-country global sourcing model with heterogeneous firms, in which firms self-select into importing based on their productivity and country-specific variables (wages, trade costs, and technology). The model delivers a simple closed-form solution for firm profits as a function of the number and characteristics of the set of countries from which a firm has invested in being able to import. A key feature of this derived profit function is that the marginal increase in profits from adding a country to the firm's set of potential sourcing locations depends on the number and characteristics of other countries in the set. This makes the analysis of the extensive margin of sourcing more complicated than in models of exporting, where entry is typically assumed to be independent across markets. Under plausible parametric restrictions, however, selection into importing features complementarity across markets. In this case, we can use standard monotone comparative statics techniques to show that the sourcing strategies of firms follow a strict hierarchical structure, as in exporting models. In our empirical implementation of the model, we also exploit these complementarities to develop an algorithm, similar to Jia (2008), to feasibly estimate the fixed costs of sourcing from different countries.

*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 are grateful to Andy Bernard, Paul Grieco, Elhanan Helpman, Marc Melitz, and Eduardo Morales for useful conversations and to BooKang Seol and Linh Vu for research assistance.

1 Introduction

In recent years, theoretical research in international trade has adopted a decidedly granular approach. The workhorse models in the literature now derive aggregate trade and multinational activity flows by aggregating the individual decisions of firms in the world economy. This novel approach is empirically anchored in a series of studies in the 1990s that demonstrate the existence of a large degree of intraindustry heterogeneity in revenue, productivity, factor inputs, and export participation across firms. With regards to export behavior, the studies show that only a small fraction of firms, those appearing to be particularly productive, engage in exporting, and that most exporting firms sell only to a few markets. This so-called extensive margin of trade has been shown to be important in understanding variation in aggregate exports across destination markets.

The typical model in this literature focuses on the export decisions of heterogeneous firms producing differentiated final goods that are demanded by consumers worldwide. In the real world, however, a significant share of the volume of international trade – possibly up to two-thirds – is accounted for by shipments of intermediate inputs (see Johnson and Noguera, 2012). Furthermore, even when focusing on finished products, the role of large retail and wholesale intermediaries in bringing consumer goods into the U.S. and other advanced economies has featured prominently in the literature (for example, see Bernard et al., 2010). As a result, and to the extent that the participation of firms into importing also varies systematically across firms, it seems plausible to posit that aggregate trade flows are shaped just as much by the aggregation of firm-level import decisions as by the aggregation of firm-level export decisions.

The previous claim might generate two natural reactions. First, and given that every international trade transaction involves an exporter and an importer, why should one care about whether the extensive margin of trade is shaped by the export or import decisions of firms? Or, in other words, is the export versus import distinction relevant for the aggregate implications of models with intraindustry heterogeneity? Second, is there any systematic evidence concerning the behavior of the extensive margin of imports that might help discipline theoretical frameworks of the global sourcing decisions of firms?

The aim of this paper is to provide answers to these questions. We begin by unveiling a series of systematic patterns in the intensive and extensive margins of U.S. imports that suggest that selection into importing is just as important in explaining aggregate imports as selection into exporting is in explaining aggregate exports. Next, we develop a multi-country general equilibrium model of global sourcing that illustrates several important differential features associated with the determination of the extensive margin of imports. Finally, we illustrate the quantitative relevance of these differences by structurally estimating the model and performing a few counterfactual exercises.

Our empirical analysis makes use of firm-level import and production data from the U.S. Census Bureau for the year 2007. We match detailed information on firm-level imports by country of origin with production and sales data for firms with at least one manufacturing establishment. In section 2, we begin by replicating the analyses in Bernard et al. (2007, 2009) to show that the extensive margins of trade (the number of firms and the number of imported products) are

important in explaining aggregate trade flows and vary systematically with firm characteristics. In a decomposition of aggregate U.S. imports by country of origin into extensive and intensive margin components, the extensive margins jointly account for 65 percent of the observed variation. We also study how these margins are shaped by standard gravity variables, such as distance and common language. Both of these variables are important factors in explaining variation in aggregate U.S. imports across countries, but variation in the intensive margin of firms' imports accounts for only a quarter of their estimated relationships. We also confirm that importing firms are larger and more productive than non-importers, and extend this known fact to show that importer premia are present even five years prior to firms starting to import.

We next show that importer premia are increasing in the number of countries from which firms source, which suggests that more productive firms might self-select into sourcing from a larger number of countries. Although the extensive margin of imports is active at both the firm and product levels, interestingly we document that it is relatively rare for U.S. firms to purchase the same product (as defined by Harmonized Schedule ten-digit code) from more than one country. For instance, more than half of the firms in our sample do not import any product from more than one country, and for 95 percent of firms, the median number of sources per product is exactly 1.00. We interpret this finding as casting doubt on the empirical relevance of frameworks in which inputs are differentiated by country of origin, as in Armington-style models.

At the end of section 2 we also provide an 'anatomy' of U.S. imports along the lines of Eaton et al. (2011). In particular, we document the existence of a hierarchical structure in the attractiveness of different countries as sources of inputs for U.S. firms. For the top ten import locations based on the number of importing firms, we count the number of firms that import from the top destination and no other countries, the number that import from the top two destinations and no other, and so so. In total, 36 percent of firms follow this pecking order, compared to only 19.6 percent that would be predicted assuming independent sourcing probabilities based on the observed frequencies in the data. These patterns are highly suggestive of firms selecting not only into importing status but also into specific importing countries. Firms that are profitable enough to jump the hurdle to a less popular source country are also more likely to import from the more popular source countries. These stylized facts motivate the development of a model in which relatively productive firms self-select into importing from particular markets based on their profitability as well as country-specific variables.

Our theoretical framework, which we describe in section 3, is a quantifiable multi-country sourcing model with heterogeneous firms in the spirit of Antràs and Helpman (2004), but without incomplete contracting or a vertical integration decision.[1] In order to tractably handle multiple sourcing countries for a firm's importing decisions, we follow Tintelnot (2012) and embed the stochastic specification of technology of Eaton and Kortum (2002) inside each individual firm. In the model, an industry is populated by a continuum of firms each producing a differentiated final-

[1]In future work, we intend to incorporate these features into the analysis and exploit the available information on related-party imports in the U.S. Census data.

good variety by combining a continuum of intermediate inputs. The only factor of production is labor. The productivity with which inputs are combined is heterogeneous across firms, as in Melitz (2003). Firms can, in principle, buy intermediate inputs from any country in the world. Nevertheless, adding a country to the set of countries a firm is able to import from requires incurring a market-specific fixed cost. As a result, relatively unproductive firms naturally opt out of importing from certain countries that are not particularly attractive sources of inputs.

Once a firm has determined the set of countries from which it has secured the ability to source inputs – which we refer to as its *global sourcing strategy* – it then learns the various firm-specific efficiency levels with which each input can be produced in each of these 'active' countries. These efficiency levels are assumed to be drawn from an extreme-value Fréchet distribution, as in Eaton and Kortum (2002). Factoring labor costs in those countries as well as transport costs, firms then decide the optimal source for each of the inputs used in production. The model delivers a simple closed-form solution for the profits of the firm as a function of its *sourcing capability*, which in turn is a function of the set of countries from which a firm has invested in being able to import as well as of the characteristics (wages, trade costs, and average technology) of these countries. The sourcing capability of a firm is an increasing function of the number of countries from which it has gained the ability to source. Intuitively, by enlarging that set, the firm benefits from greater competition among suppliers, and thereby lowers the effective cost of its intermediate input bundle. The choice of a sourcing strategy therefore trades off lower variable cost of production against the greater fixed costs associated with a more complex global sourcing strategy. Quite intuitively, we show that more productive firms necessarily choose strategies that confer them (weakly) higher sourcing capabilities, which implies that their cost advantage is magnified by their sourcing decisions, thus generating an increased skewness in the size distribution of firms.

A key characteristic of the derived profit function is that a firm's marginal gain in profits associated with adding a country to its set of sourcing countries generally depends on the number and characteristics of the other countries in this set. This contrasts with standard models of selection into exporting featuring constant marginal costs, in which the decision to service a given market is independent of that same decision in other markets. Whether the decisions to source from different countries are complements or substitutes crucially depends on a parametric restriction involving the elasticity of demand faced by the final-good producer and the Fréchet parameter governing the variance in the distribution of firm-specific input efficiencies across locations. Selection into importing features complementarity across markets whenever demand is relatively elastic (so profits are particular responsive to variable cost reductions) and whenever input efficiency levels are relatively heterogeneous across markets (so that the reduction in expected costs achieved by adding an extra country in the set of active locations is relatively high).

Conversely, when demand is inelastic or input efficiency draws are fairly homogeneous, the addition of country to a firm's global sourcing strategy instead reduces the marginal gain from adding other locations. In such a situation, the problem of a firm optimally choosing its sourcing strategy is extremely hard to characterize, both analytically as well as quantitatively, since it boils

down to solving a combinatorial problem with 2^J elements (where J is the number of countries) with little guidance from the model.

The case with complementary sourcing decisions turns out to be much more tractable and delivers sharp results reminiscent of the broad patterns discussed in section 2. In particular, we can use standard tools from the monotone comparative statics literature to show that the sourcing strategies of firms follow a strict hierarchical structure in which the number of countries in a firm's sourcing strategy is (weakly) increasing in the firm's core productivity level. Furthermore, we can show that the number of firms sourcing from a particular country is (weakly) increasing in the cost attractiveness of that location, with that attractiveness being shaped by the wages, technology, transport costs and fixed costs of offshoring associated with that sourcing location.

We next outline how the aggregation of firms' sourcing decisions shapes aggregate input flows across countries. As in heterogeneous firms models of exporting, the extensive margin of global sourcing amplifies the response of trade flows to changes in variable trade costs. In fact, in the knife-edge case in which parameter values are such that the sourcing decisions are independent across markets, our model delivers a gravity equation for input flows that is identical to those derived by Chaney (2008), Arkolakis et al. (2008), or Helpman et al. (2008) in models of exporting, with the trade elasticity being uniquely pinned down by the shape parameter of the Pareto distribution characterizing the dispersion in core productivity across final-good producers. When we depart from that knife-edge case, however, our model gives rise to an *extended* gravity equation featuring third market effects, which invalidates the estimation of the aggregate trade elasticity with standard approaches. For similar reasons, the interdependencies in the global sourcing decisions of firms undermine the usefulness of the sufficient statistic approach to welfare analysis advocated by Arkolakis et al. (2012).

Our quantitative analysis enables separate identification of the sourcing potential (a function of technology, trade cost, and wages) of a country and the fixed cost of sourcing from that country. In a first step, we use the structure of the model to recover the sourcing potential of 64 foreign countries from U.S. firm-level data on the intensive margin of intermediate input purchases. In the second stage, we use the estimated sourcing potentials and data on average markups in U.S. manufacturing to estimate the elasticity of final-good demand and the Fréchet parameter governing the dispersion in productivity of intermediate good producers. We find robust evidence suggesting that the extensive margin sourcing decisions of U.S. firms are complements, which provides a rationale for the systematic patterns documented in section 2. Furthermore, our estimates imply that a firm which sources from all foreign countries faces 7 to 11 percent lower variable cost, with the exact value depending on the preferred specification for the estimation of the dispersion parameter.[2] In our third and final step, we estimate the fixed costs of sourcing associated with different countries by using a simulated method of moments. In doing so, we adopt the iterative algorithm developed by Jia (2008), which exploits the complementarities in the 'entry' decisions of

[2]Quantitatively, this is comparable to the findings of Halpern et al. (2011) for Hungarian firms. Using a two-country model and a method similar to Olley and Pakes (1996), they find that importing all foreign varieties would increase firm productivity of a Hungarian firm by 12 percent.

firms, and uses lattice theory to greatly reduce the dimensionality of the firm's optimal sourcing strategy problem. We estimate median fixed cost of sourcing from a country of 40-60 thousand U.S. dollars and find around 30 percent lower fixed cost of sourcing for countries with a common language. Our estimated model matches moments of the size distributions of U.S. firms' input purchases and the shares of firms importing from foreign countries. We are in the process of using our estimates from the model to perform a series of counterfactual exercises.

Our paper is related to several strands of the literature. As mentioned above, we build on the approach in Tintelnot (2012) of applying the Eaton and Kortum (2002) stochastic representation of technology to the problem of a firm optimally choosing a production location in a multi-country world. The focus in our paper is nevertheless quite distinct. While Tintelnot (2012) studies the location of final-good production of multi-product multinational firms, we instead develop a model of global sourcing with multiple capability sources of inputs. Rodríguez-Clare (2010) and Garetto (2013) also adapt the Eaton and Kortum (2002) framework to the study of global sourcing decisions but in two-country models and with very different goals in mind.[3]

Several recent papers have explored the implications of changes in the extensive margin of intermediate input imports on firm performance and aggregate productivity. These include the studies by Amiti and Konings (2007) for Indonesia, Goldberg et al. (2010) and Loecker et al. (2012) for India, Halpern et al. (2011) for Hungary, and Gopinath and Neiman (2013) for Argentina. Because the focus of these papers is largely empirical, these authors develop simple two-country models in which domestic and foreign inputs are assumed to be imperfectly substitutable, so that the productivity improvements resulting from an increase in imports of intermediate inputs are associated with love-for-variety effects, as in Ethier (1982). Instead, in our framework, inputs are not differentiated by country of origin, and the gains in profitability associated with an expansion in the extensive margin of imports are brought about by an increase in competition across input sources. Moreover, our model can flexibly accommodate an arbitrary number of countries and is thus a more reliable tool for quantitative and counterfactual analysis.

Our paper is not the first one to describe the inherent difficulties in solving for the extensive margin of imports in a model with multiple intermediate inputs. Blaum et al. (2013) discuss the existence of interdependencies across sourcing decisions analogous to those in our model (although with a love-for-variety model in mind), and conclude that their model provides no general predictions for the extensive margin of importing. Instead, we show that under certain plausible parametric restrictions, the problem can actually be characterized with standard optimization techniques. [4]

The focus in our paper is on international trade transactions in which firms are both on the buying and selling side of the transaction. As such, our work is to some extent related to a

[3]See also Ramondo and Rodríguez-Clare (2013) and Arkolakis et al. (2013) for recent related contributions to quantitative models of multinational activity.

[4]The seminal applications of the mathematics of complementarity in the economics literature are Vives (1990) and Milgrom and Roberts (1990). Grossman and Maggi (2000) and Costinot (2009) are particularly influential applications of these techniques in international trade environments.

burgeoning literature in international trade focused on firm-to-firm that includes the work of Blum et al. (2010), Eaton et al. (2012), Eaton et al. (2013), Carballo et al. (2013), Bernard et al. (2013), and Dragusanu (2014).

Finally, at a broader level, our paper naturally relates to the vast literature on offshoring, and particularly to the work of Antràs et al. (2006), Grossman and Rossi-Hansberg (2008, 2012), and Fort (2013) who all emphasize the importance of heterogeneity for understanding the margins of offshoring. Relative to those papers, our focus is on the role of heterogeneous productivity (rather than heterogeneous skills or offshoring costs), and we also provide a tractable multi-country model, while those papers develop simpler two-country frameworks.

The rest of the paper is structured as follows. We begin in section 2 by characterizing a few salient features of the extensive and intensive margins of U.S. imports. In section 3, we develop our multi-country model of global sourcing, which we then structurally estimate in section 4. In section 5, we (will) perform a few counterfactual exercises, and finally we (will) provide some concluding comments in section 6.

2 The Margins of U.S. Imports: Stylized Facts

In this section, we use data on U.S. firms' production, sales, and import behavior to document a number of key facts about the margins of firms' import behavior.

2.1 Data Description

The primary data used in the paper are from the U.S. Census Bureau's 2007 Economic Censuses (EC), Longitudinal Business Database (LBD), and Import transaction database. The LBD uses administrative record data to provide employment and industry for every private, non-farm, employer establishment in the U.S. The ECs supplement this information with additional establishment-level variables, such as sales, value-added, and input usage. While there is extensive research that uses the Census of Manufactures (CM), our study is one of the few to include information from all Economic Censuses.[5] This coverage ensures that we provide an accurate depiction of the entire firm. For example, a firm engaged in offshoring may have less domestic employment and sales in manufacturing, but considerably more in wholesale and services. The import data, collected by U.S. Customs facilities, are based on the universe of import transactions into the U.S. They contain information on the products, values, countries, and related party status of firms' imports. We match these data at the firm level to LBD and the EC data.

The focus of this paper is on firms involved in the production of goods. We therefore limit the analysis to firms with at least one manufacturing establishment. Because we envision a production process entailing physical transformation activities (manufacturing) as well as headquarter

[5]The other censuses are for Construction, Finance, Insurance and Real Estate, Management of Companies, Professional and Technical Services, Retail Trade, Transportation and Warehousing, and Wholesale Trade. The variables available differ across these censuses.

activities (design, distribution, marketing, etc.), we include firms with activities outside of manufacturing.[6] We also limit the sample to firms with positive sales and employment and exclude all mineral imports from the analysis since they do not represent offshoring. Firms with at least one manufacturing plant account for five percent of firms, 23 percent of employment, 38 percent of sales, and 65 percent of non-mineral imports. In terms of explaining aggregate U.S. sourcing patterns, it is critically important to include firms with manufacturing and other activities. They account for 60 percent of U.S. imports, while manufacturing-only firms account for just five percent. The import behavior of the firms in our sample is consistent with patterns in past work on heterogeneous firms in trade. Only one quarter of U.S. manufacturing firms have positive imports in 2007. Additional details on the sample and data construction are in the Data Appendix.

2.2 Stylized Facts

We use this comprehensive dataset to document a number of facts about U.S. firms' import behavior. First, we confirm the importance of the extensive margin in our sample, both in explaining variation in aggregate trade flows across countries, and in accounting for how these flows are systematically related to standard gravity variables. Next, we provide evidence strongly suggestive of selection into importing. Importing firms are larger and more productive than non-importers, and these import premia are present before firms begin to import. Extending the information on premia to a multi-country setting, we provide new evidence that importer premia are increasing in the number of countries from which firms source. We also provide new facts on the number of countries from which a firm sources the same product. Finally, we shed new light on the extent to which firms follow a hierarchy pattern in terms of the countries from which they source.

2.2.1 Decomposition of Intensive and Extensive Margins

We begin by confirming the importance of the extensive margins of trade, both in terms of the number of imported products and the number of importing firms, first documented by Bernard et al. (2009). Following those authors, we decompose total U.S. imports $M_{US,j}$ from country j according to

$$ln(M_{US,j}) = ln(N^{firms}_{US,j}) + ln(N^{prods}_{US,j}) + ln\left(\frac{O_{US,j}}{N^{firms}_{US,j} \times N^{prods}_{US,j}}\right) + ln\left(\frac{M_{US,j}}{O_{US,j}}\right),$$

where $O_{US,j}$ is the number of firm-product combinations with positive imports from j. The first two terms represent the unique numbers of firms ($N^{firms}_{US,j}$) importing and products ($N^{prods}_{US,j}$) imported from country j. The third term, referred to as the density, captures the fraction of firm-production

[6] We recognize that focusing on firms with positive manufacturing activity will miss some offshoring, for example by factoryless goods producers (FGPs) in the wholesale sector that have offshored all physical transformation activities (see Bernard and Fort, 2013, for details). Unfortunately, there is no practical way to distinguish all FGPs from traditional wholesale establishments. Furthermore, data on value-added and input usage, which is crucial for our estimation in section 4, is scant for firms outside manufacturing is sparse. It should also be noted that we cannot identify manufacturing firms that use inputs imported by intermediaries.

combinations with positive import values. The final term captures the intensive margin. It measures the average import value per firm-product observation, for all combinations with positive imports. Table 1 presents coefficients from OLS regressions of the logarithm of each margin on the logarithm of total trade. As is well known, these OLS coefficients sum to one, with each coefficient representing the share of overall variation explained by each margin. As in previous work, we find that variation in the extensive margins account for the majority of the variation in aggregate import volume across countries. The extensive margins account for a total of 65 percent, while the intensive margin explains just 35 percent of the total variation.

Table 1: Extensive and Intensive Margin Decomposition

	Log of number of importing firms	Log of number of imported products	Log of Density	Log of average import value per product per firm
	0.541***	0.535***	-0.426***	0.350***
	(0.016)	(0.015)	(0.014)	(0.018)
Adj. R^2	0.85	0.84	0.81	0.64
Observations	221	221	221	221

Notes: Each column corresponds to results from regressing the log of each margin on the log of total import values. The coefficients are a measure of the fraction of variation in aggregate import volumes across countries explained by that margin. Density represents the fraction of all possible firm-product combinations with positive import values. The estimated coefficients sum to one.

A similar decomposition can be performed to assess the extent to which the extensive and intensive margins are systematically related to standard gravity variables. To do so, we separately regress the log of total imports and the log of each margin on the log of country GDP, log of distance, and an indicator for common language.[7] In this exercise, the OLS coefficients on the margins must sum to the coefficient for total imports.

Table 2 presents the results. The first column shows the well-known result that aggregate imports are increasing in country GDP and common language, but decreasing in distance. The more interesting findings from the decomposition are evident in comparisons of the importance of these standard gravity variables across margins. Distance and common language have large and statistically significant relationships with the numbers of firms and products imported across countries. A one percent increase in distance is associated with a 1.1 percent decrease in the number of importers and the number of imported products. In contrast, a one percent increase in distance is only associated with a 0.4 percent decrease in the average import value per product per firm, and this effect is only statistically significant at the ten percent level. Common language has

[7]This decomposition is similar to one presented in Bernard et al. (2007). The analysis here differs by including the density measure and a common language indicator. We note that the estimated effect of distance is significantly larger in our analysis.

no statistically significant relationship with the intensive margin. These results suggest that the gravity variables that have been so successful at explaining aggregate trade flows work primarily through the extensive margins of trade.

Table 2: Gravity Decomposition

	Log of total import value	Log of number of importing firms	Log of number of imported products	Density	Log of average import value per product per firm
Log of GDP	1.373***	0.872***	0.829***	-0.666***	0.339***
	(0.058)	(0.035)	(0.038)	(0.030)	(0.044)
Log of Distance	-1.651***	-1.057***	-1.079***	0.900***	-0.415*
	(0.294)	(0.176)	(0.191)	(0.153)	(0.222)
Common Language	0.961***	0.709***	0.501***	-0.494***	0.246
	(0.296)	(0.178)	(0.192)	(0.154)	(0.224)
Adj. R^2	0.78	0.79	0.75	0.75	0.26
Observations	176	176	176	176	176

Notes: Notes: Each column corresponds to a separate regression of the log of that margin on the gravity variables. Density represents the fraction of all possible firm-product combinations with positive import values. The estimated coefficients in columns 2-5 sum to the coefficient in column one.

2.2.2 Selection into Importing

Having documented the importance of the extensive margins, we now show that importers differ systematically from non-importers. Following Bernard et al. (2007), we report employment, sales, and productivity premia for firms that import in 2007. To do so, we regress the log each of these variables on an importer dummy and industry controls (see the Data Appendix for details). Table 3 reports the results. The top panel presents results using 2007 values of firm size and productivity and the bottom panel uses 2002 values. The first column of the table shows that firms importing in 2007 are larger and more productive than non-importers. In addition, these premia for 2007 import status were present in 2002. The magnitude of these import premia is similar to those typically found for exporters, with importers being on average about three times larger and about 6-7% more productive than non-importers.

While these importer premia are consistent with selection into importing, they may simply be a result of a firm's import activities. For example, Amiti and Konings (2007) show that importing leads to increased productivity at the firm level. To determine whether the premia existed prior to importing, the right column of Table 3 reports the same premia for the subset of firms that did not import in 2002. The results here support the premise of selection into importing. Firms that began importing after 2002 were already larger and more productive in 2002.

Finally, we assess the extent to which firm size and productivity relate to the number of countries

Table 3: Premia for 2007 Importers

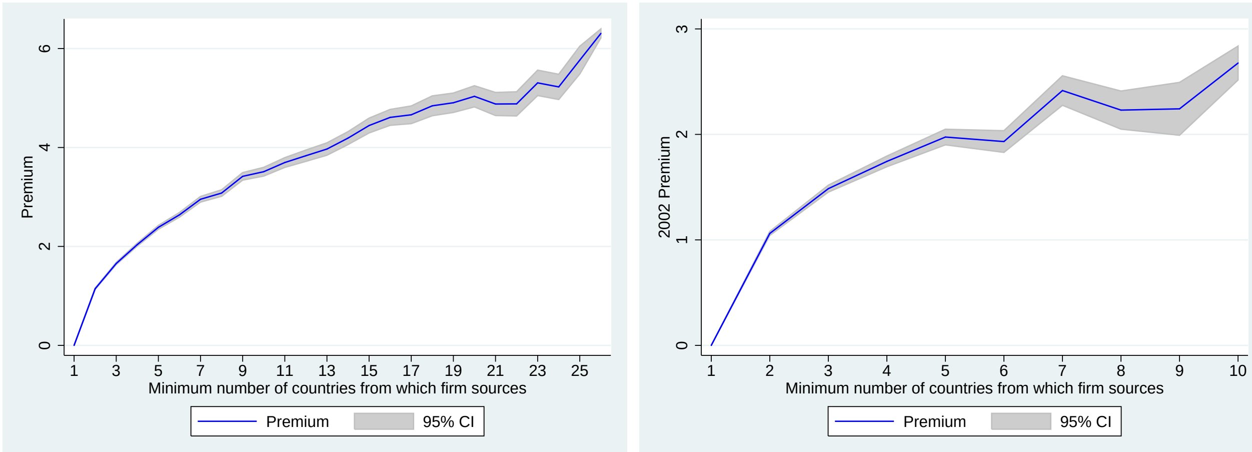
	All Firms	Non-2002 Importers
2007 Log employment	1.552***	1.269***
2007 Log sales	1.737***	1.399***
2007 Log value-added per worker	0.060***	0.039***
2002 Log employment	1.466***	1.154***
2002 Log sales	1.638***	1.270***
2002 Log value-added per worker	0.074***	0.052***

Notes: All results are from OLS regressions of the variable listed on the left on an indicator equal to one if the firm imported in 2007. The first column includes all firms. The second column is based on the subset of firms that did not import in 2002. Results with 2002 variables are based only on the subset of firms that existed in 2002. All regressions include four digit industry controls.

from which a firm sources. Figure 1a plots sales premia by firms' minimum number of sourcing countries. To construct the figure, we regress the log of firm sales on cumulative dummies for the number of countries from which a firm sources. The omitted category is non-importers, so the premia are interpreted as the difference in size between non-importers and firms that import from at least one country, at least two countries, etc. The horizontal axis denotes the number of countries from which a firm sources, with 1 corresponding to firms that use only domestic inputs. Firms that import from 25 or more countries are grouped together as one category due to confidentiality protection rules. We plot the sum of these coefficients along with the 95 percent confidence interval. There is a strong, positive, statistically significant relationship between the number countries from which a firm sources and its sales. Firms that import from one country are about twice the size of non-importers, firms that source from 13 are about four times as large, and firms sourcing from 25 or more countries are over six times the size of non-importers. Somewhat remarkably, each estimated coefficient is positive until about twenty countries. The same analysis with employment and productivity yields similar patterns.

To ensure that the increasing relationship between size and the number of sourcing countries is not a result of firms' offshoring decisions, we again restrict the sample to firms that did not import in 2002. Although firms that did not import in 2002 generally source from fewer countries in 2007, the same pattern emerges. Figure 1b plots the estimated coefficients and confidence intervals, this time grouping all firms that import from ten or more countries together. As for the full sample, an increase in the number of source countries is strongly associated with an increase in firm size.

Figure 1: Sales premia and minimum number of sourcing countries in 2007



(a) 2007 Sales premia

(b) 2002 Sales premia for non-2002 importers

Notes: Figures plot the sum of coefficients and 95 percent confidence interval from regressions of the log of sales on cumulative dummies for the number of countries from which a firm sources in 2007. The omitted category is firms that source only domestically. Panel (a) plots 2007 sales premia. Panel (b) is based only on firms that did *not* import in 2002 and plots 2002 sales premia.

2.2.3 The Extensive Margin at the Product Level

The results in Tables 1 and 2 indicate that the extensive margin of imports is active at both the firm and product levels. We next document, however, that when firms source from multiple countries, they seldom buy the same product from more than one country. To do so, we begin by defining a product as a distinct Harmonized Schedule ten-digit code, of which there are nearly 17,000 categories in U.S. data. We next use data on all available firm-product combinations to calculate the mean, median, and 95th percentile number of countries per product in this firm-product-level data. Column 1 in Table 4 shows that the average number of countries per product for a firm is 1.47, the median is 1 and the 95th percentile is 3.[8]

A small fraction of extremely large firms imports many more products, and from many more countries, than most firms. To gain a clear understanding of how the number of HS10 products per country varies across all firms, we aggregate the firm-product-level data to the firm level, calculating the mean, median, and maximum number of countries per imported product for each firm. For a firm importing a single product, these three numbers naturally coincide. Columns 2-4 present the mean, median, and 95th percentiles of these firm-level measures. The median firm imports a single product from an average of only 1.03 countries. The 95th percentile of the firm mean is below two, with a mean of just 1.78. The median number of countries per product for firms is always 1.00, even for the 95th percentile of firms. Finally, the maximum number of countries per product for the median firm is still just 1.00, while firms in the 95th percentile import the same product from

[8]Data confidentiality protection rules preclude us from disclosing exact percentiles. Statistics for all percentiles are therefore the average number of countries per product for all firms within +/- one percent of a given percentile.

Table 4: Firm-level statistics on the number of countries per HS10 product

	Firm-HS Level	Firm Level Mean	Median	Max
Mean	1.47	1.11	1.00	1.61
Median	1.00	1.03	1.00	1.00
95%tile	3.32	1.78	1.00	4.36

Notes: Column 1 reports statistics from firm-product-level data on the number of countries from which a firm imports the same HS10 product. Columns 2-4 report statistics from firm-level data of the firm mean, firm median and firm maximum of the number of countries from which a firm imports the same HS10 product.

a maximum of about 4 countries.

These findings suggest that a framework in which inputs are simply differentiated by country of origin, as in Armington-style models, cannot explain a significant portion of firms' sourcing decisions. Blaum et al. (2013) perform a similar exercise with French firm-level data and report that 80 percent of French firms import the same product from less than three countries, while the top one percent of firms import the same product from more than eight countries. A key difference between their exercise and ours is that they define a product as an eight-digit code, and that lower level of disaggregation naturally leads them to obtain higher measures of importing countries per product.[9]

2.2.4 Systematic Patterns in the Extensive Margin of Importing

The facts presented thus far highlight the importance of firm selection, and the extensive margin more generally, in explaining U.S. imports. We now provide evidence reminiscent of Eaton et al. (2011) of systematic patterns in this extensive margin.

As a preliminary step, Table 5 lists the top 10 sourcing countries for U.S. manufacturers in 2007, based on the number of importing firms. These countries account for 93 percent of importers in our sample and 74 percent of imports. The first two columns provide the country rank by number of firms and by import value. Canada ranks number one for manufacturers in both dimensions. There are significant differences in these ranks, however, for other countries. China is number two for firms but only number three for value. Mexico, the number two country in terms of value, ranks eighth in terms of the number of importing firms. Columns 3-4 provide details on the number of firms and their fraction of total importers and columns 5-6 give similar information for import values. It is clear that there are significant differences across countries in extensive and intensive margins. For

[9] We also note that, even at the HS10 level, certain products such as those described as "Services" or "NESOI" are likely to consist of multiple inputs. In addition, large multi-unit firms may source the same input from different countries in order to minimize transport costs. Fort (2013) finds that heterogeneity in U.S. firms' domestic distance to foreign suppliers is an important predictor of their extensive margin sourcing decisions.

example, the U.K. and Taiwan account for only three and two percent of total imports respectively, but 18 percent of all importers source from the U.K. and 16 percent source from Taiwan. These considerable divergences between the intensive and extensive margins of trade are suggestive of the importance of heterogeneity in the fixed costs of sourcing from particular countries. Under this interpretation, Table 5 appears to indicate that relative to the sourcing potential of these countries, fixed costs of sourcing are disproportionately high in Mexico and Japan and low in Italy and Taiwan. In section 4, we will employ a similar approach, although augmented to account for richer moments of the data, to estimate country-specific fixed costs of sourcing.

Table 5: Top 10 source countries for U.S. firms, by number of firms

	Rank by:		Number of Importers		Value of Imports	
	Firms	Value	Firms	% of Total	Imports	% of Total
Canada	1	1	37,800	59	145,700	16
China	2	3	21,400	33	121,980	13
Germany	3	5	13,000	20	62,930	7
United Kingdom	4	6	11,500	18	30,750	3
Taiwan	5	11	10,500	16	16,630	2
Italy	6	13	8,500	13	13,230	1
Japan	7	4	8,000	12	112,250	12
Mexico	8	2	7,800	12	125,960	14
France	9	9	6,100	9	22,980	3
Korea, South	10	10	5,600	9	20,390	2

Notes: Number of firms rounded to nearest 100 for disclosure avoidance. Imports in millions of $s, rounded to nearest 10 million for disclosure avoidance. This is an alternative version of the table above.

Focusing again on the extensive margin of importing, we next assess the extent to which firms follow a hierarchical pecking order in their import behavior, as found by Eaton et al. (2011) in their study of French exporters. Specifically, we count the number of firms that import from Canada (the top destination by firm rank) and no other countries, the number that import from Canada and China (the top two destinations) and no others, the number that import from Canada, China, and Germany and no others, and so on. When calculating these numbers, we limit the analysis to the top ten countries by firm rank. Table 6 presents the results. The first column shows the number of firms that import from each string of countries. To assess the significance of these numbers, column three provides the number of firms that would import from each string if sourcing decisions were independent and the fraction of actual firms importing from a given country represented that independent probability. The total fraction of firms that follows a pecking order is 36.0, almost twice the 19.6 percent that would be observed under independence. These patterns are highly suggestive of a pecking order in which country characteristics make some countries particularly appealing for all U.S. firms.

Table 6: U.S. firms importing from strings of top 10 countries

String	Data		Under Independence	
	Firms	% of Importers	Firms	% of Importers
CA	17,980	29.82	6,760	11.21
CA-CH	2,210	3.67	3,730	6.19
CA-CH-DE	340	0.56	1,030	1.71
CA-CH-DE-GB	150	0.25	240	0.40
CA-CH-DE-GB-TW	80	0.13	50	0.08
CA-CH-DE-GB-TW-IT	30	0.05	10	0.02
CA-CH-DE-GB-TW-IT-JP	30	0.05	0	0.00
CA-CH-DE-GB-TW-IT-JP-MX	50	0.08	0	0.00
CA-CH-DE-GB-TW-IT-JP-MX-FR	160	0.27	0	0.00
CA-CH-DE-GB-TW-IT-JP-MX-FR-KR	650	1.08	0	0.00
TOTAL Following Pecking Order	21,680	36.0	11,820	19.6

Notes: The string CA means importing from Canada but no other among the top 10; CA-CH means importing from Canada and China but no other, and so forth. % of Importers shows percent of each category relative to all firms that import from top 10 countries.

3 Theoretical Framework

Guided by these empirical regularities, in this section we develop a quantifiable multi-country model of global sourcing.

3.1 Preferences and Endowments

Consider a world consisting of J countries where individuals value the consumption of differentiated varieties of manufactured goods according to a standard symmetric CES aggregator

$$U_M = \left(\int_{\omega \in \Omega_j} q(\omega)^{(\sigma-1)/\sigma} \, d\omega \right)^{(\sigma/\sigma-1)}, \quad \sigma > 1, \tag{1}$$

where Ω_j is the set of manufacturing varieties available to consumers in country $j \in J$ (with some abuse of notation we denote by J both the number as well as the set of countries). These preferences are assumed to be common worldwide and give rise to the following demand for variety ω in country j:

$$q_j(\omega) = E_j P_j^{\sigma-1} p_j(\omega)^{-\sigma}, \tag{2}$$

where $p_j(\omega)$ is the price of variety ω, P_j is the standard ideal price index associated with (1), and E_j is aggregate spending on manufacturing goods in country j. For what follows it will be useful

to define a (manufacturing) market potential term for market j as follows

$$B_j = \frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} E_j P_j^{\sigma-1}. \tag{3}$$

There is a unique factor of production, labor, which commands a wage w_j in country j. We will discuss below two alternative ways to close the model in general equilibrium. In the first one, we think of the manufacturing sector as being the only sector of the economy and let the amount of labor L_j available to manufacturing firms be inelastically supplied. In the second one, we introduce an additional sector in the economy that also employs labor and that is large enough to pin down wages in terms of that 'outside' sector's output.

3.2 Technology and Market Structure

There exists a measure N_j of final-good producers in each country $j \in J$ and each of these producers owns a blueprint to produce a single differentiated variety. The market structure of final good production is characterized by monopolistic competition. We will discuss below the cases in which N_j is exogenously predetermined and in which it is pinned down by a free entry condition.

Production of final-good varieties requires the assembly of a bundle of *intermediates*. Intermediates can be offshored and a key feature of the equilibrium will be determining the location of production of different intermediates. The bundle of intermediates contains a continuum of measure one of inputs, assumed to be imperfectly substitutable with each other, with a constant and symmetric elasticity of substitution equal to $1/(1-\rho)$. Very little will depend on the particular value of ρ.

All intermediates are produced with labor under *firm-specific* technologies featuring constant returns to scale. We will index final-good firms by their 'core productivity', which we denote by φ, and which governs the mapping between the bundle of inputs and final-good production. We will denote by $a_j(v,\varphi)$ the firm-specific unit labor requirements associated with firm φ procuring intermediate $v \in [0,1]$ in country $j \in J$. Below, we will specify in more detail the structure of these intermediate costs.

With the above assumptions and notation at hand, we can express the marginal cost for firm φ based in country i of producing a unit of a final-good variety as

$$c_i\left(\{j(v)\}_{v=0}^{1}, \varphi\right) = \frac{1}{\varphi}\left(\int_0^1 \left(\tau_{ij(v)} a_{j(v)}(v,\varphi)\, w_{j(v)}\right)^{1-\rho} dv\right)^{1/(1-\rho)}, \tag{4}$$

where $\{j(v)\}_{v=0}^{1}$ corresponds to the infinitely-dimensional vector of locations of intermediate input production and $\tau_{ij(v)}$ denotes the iceberg trade costs between the base country i and the production location $j(v)$. For simplicity, we assume that final-good varieties are prohibitively costly to trade across borders, although we will later relax this assumption and briefly study the joint determination of the extensive margins of both exports and imports.

Final-good producers are heterogeneous in their core productivity level φ. Following Melitz (2003), we assume that firms draw their value of φ from a country-specific distribution $g_i(\varphi)$, with support in $[\underline{\varphi}_i, \infty)$, and with an associated continuous cumulative distribution $G_i(\varphi)$.

Building on Eaton and Kortum (2002), we treat the (infinite-dimensional) vector of firm-specific intermediate input efficiencies $a_j(v, \varphi)$ as the realization of an extreme value distribution. More specifically, firm φ draws the value of $a_j(v, \varphi)$ for a given location j from the Fréchet distribution

$$\Pr(a_j(v, \varphi) \leq a) = e^{-T_j a^{-\theta}}, \quad \text{with } T_j > 0. \tag{5}$$

These firm-specific draws are assumed independent across locations and inputs. Note that we specify these draws as being independent of the firm's core productivity, φ, but a higher core productivity enables firms to transform the bundle of intermediates into more units of final goods. As in Eaton and Kortum (2002), T_j governs the state of technology in country j, while θ determines the variability of productivity draws across inputs, with a lower θ fostering the emergence of comparative advantage *within* the range of intermediates sector across countries.[10] It will be convenient to denote by $\mathbf{a}_j(\varphi) \equiv \{a_j(v, \varphi)\}_{v=0}^{1}$ the vector of of unit labor requirements drawn by firm φ in country j.

Although firms can potentially draw a vector $\mathbf{a}_j(\varphi)$ for each country $j \in J$, we assume that a firm from country i only acquires the capability to offshore intermediates to j and learns this vector after incurring a fixed cost equal to f_{ij} units of labor in country i (at a cost $w_i f_{ij}$). We denote by $\mathcal{J}_i(\varphi) \subseteq J$ the set of countries for which a firm based in i with productivity φ has paid the associated fixed cost of offshoring $w_i f_{ij}$. For brevity, we will often refer to $\mathcal{J}_i(\varphi)$ as the *sourcing strategy* of that firm.

Apart from these fixed costs of offshoring, whenever we consider the equilibrium with free entry of final-good producers, we will also incorporate a cost of entry equal to f_{ei} units of labor in the country where the bundle of intermediate goods is assembled into the final good (country i in our notation above). In such a case, producers only learn their productivity φ after paying the entry cost, but are assumed to choose their sourcing strategy with knowledge of that core productivity level. We will sometimes refer to workers employed in input production as *production workers* and to those employed to cover fixed costs as *non-production workers*. This will facilitate the structural estimation of model in section 4, but is immaterial for the equilibrium, since these different occupations provide the same remuneration for workers.

This completes the description of the key assumptions of the model. We will solve for the equilibrium of the model in four steps. First, we describe optimal firm behavior conditional on a given sourcing strategy $\mathcal{J}_i(\varphi)$. Second, we characterize the choice of this sourcing strategy and relate our results to the stylized facts unveiled in section 2. Third, we aggregate the firm-level decisions and discuss alternative ways to close the model in general equilibrium. Finally, we discuss special cases and extensions of our framework.

[10] For technical reasons, we assume $\theta > \rho/(1-\rho)$. Apart from satisfying this restriction, the value of ρ does not matter for any outcomes of interest and will be absorbed into a constant.

3.3 Firm Behavior Conditional on a Sourcing Strategy

Consider a firm based in country i with productivity φ that has incurred all fixed costs associated with a given sourcing strategy $\mathcal{J}_i(\varphi)$. In light of the cost function in (4), it is clear that after drawing the vector $\mathbf{a}_j(\varphi)$ for each country $j \in \mathcal{J}_i(\varphi)$, the firm will choose the location of production of any input v that solves $\min_j \{\tau_{ij} a_j(v,\varphi) w_j\}$. Although such a problem might seem daunting to characterize at first glance, we can follow closely the analysis in Eaton and Kortum (2002) to provide a simple characterization of it. In particular, using the properties of the Fréchet distribution in (5), one can show that the firm will source a positive measure of intermediates from each country in its sourcing strategy set $\mathcal{J}_i(\varphi)$. Furthermore, the share of intermediate input purchases sourced from any country j (including the home country i) is simply given by

$$\chi_{ij}(\varphi) = \frac{T_j (\tau_{ij} w_j)^{-\theta}}{\Theta_i(\varphi)} \quad \text{if } j \in \mathcal{J}_i(\varphi) \tag{6}$$

and $\chi_{ij}(\varphi) = 0$ otherwise, where

$$\Theta_i(\varphi) \equiv \sum_{k \in \mathcal{J}_i(\varphi)} T_k (\tau_{ik} w_k)^{-\theta}. \tag{7}$$

The term $\Theta_i(\varphi)$ summarizes the *sourcing capability* of firm φ from i. Note then that each country j's market share in the firm's purchases of intermediates corresponds to this country's contribution to this sourcing capability $\Theta_i(\varphi)$. Countries in the set $\mathcal{J}_i(\varphi)$ with lower wages w_j, more advanced technologies T_j, or lower distance from country i are predicted to have higher market shares in the intermediate input purchases of firms based in country i. We shall refer to the term $T_j(\tau_{ij} w_j)^{-\theta}$ as the *sourcing potential* of country j from the point of view of firms in i.

It may seem surprising that the dependence of country j's market share $\chi_{ij}(\varphi)$ on wages and trade costs is shaped by the Fréchet parameter θ and not by the substitutability across inputs, as governed by the parameter ρ in equation (4). The reason for this, as in Eaton and Kortum (2002), is that variation in market shares is explained exclusively by a *product-level* extensive margin. More specifically, firm φ from i sources a disproportionate measure of inputs from locations in its sourcing strategy $\mathcal{J}_i(\varphi)$ featuring particularly favorable production costs (as summarized by $T_j(\tau_{ij} w_j)^{-\theta}$). This oversampling in turn puts upwards pressure on the average price that firm φ from i ends up paying for inputs purchased from those attractive locations. As it turns out, these two effects on prices (low costs but oversampling) exactly cancel each other to the point at which the distribution of input prices paid by firm φ from i in any country $j \in \mathcal{J}_i(\varphi)$ is identical. As a result, variation in market shares simply reflects the fact that firms buy more inputs from more attractive sourcing locations.

After choosing the least cost source of supply for each input v, the overall marginal cost faced

by firm φ from i can be expressed, after some nontrivial derivations, as

$$c_i(\varphi) = \frac{1}{\varphi}\left(\gamma\Theta_i(\varphi)\right)^{-1/\theta}, \tag{8}$$

where $\gamma = \left[\Gamma\left(\frac{\theta+1-\rho}{\theta}\right)\right]^{\theta/(1-\rho)}$ and Γ is the gamma function.[11] Note that in light of equation (7), the addition of a new location to the set $\mathcal{J}_i(\varphi)$ increases the sourcing capability of the firm and necessarily lowers its effective marginal cost. Intuitively, an extra location grants the firm an additional cost draw for all varieties $v \in [0,1]$, and it is thus natural that this greater competition among suppliers will reduce the expected minimum sourcing cost per intermediate. In fact, the addition of a country to $\mathcal{J}_i(\varphi)$ lowers the expected price paid for *all* varieties v, and not just for those that are ultimately sourced from the country being added to $\mathcal{J}_i(\varphi)$. This feature of the model distinguishes our framework from Armington-style love-for-variety models, in which the addition of an input location also decreases costs and increases revenue-based productivity, but in which the price paid for a country's variety is unaffected by the inclusion of other countries in the set $\mathcal{J}_i(\varphi)$.[12]

Using the demand equation (2) and the derived cost function in (8), we can express the firm's profits conditional on a sourcing strategy $\mathcal{J}_i(\varphi)$ as

$$\pi_i(\varphi) = \varphi^{\sigma-1}\left(\gamma\Theta_i(\varphi)\right)^{(\sigma-1)/\theta} B_i - w_i \sum_{j\in\mathcal{J}_i(\varphi)} f_{ij}, \tag{9}$$

where B_i is given in (3). As is clear from equation (9), when deciding whether to add a new country j to the set $\mathcal{J}_i(\varphi)$, the firm trades off the reduction in costs associated with the inclusion of that country in the set $\mathcal{J}_i(\varphi)$ – which increases the sourcing capability $\Theta_i(\varphi)$ – against the payment of the additional fixed cost $w_i f_{ij}$. We next turn to studying the optimal determination of the sourcing strategy $\mathcal{J}_i(\varphi)$.

3.4 Optimal Sourcing Strategy

Each firm's optimal sourcing strategy is a combinatorial optimization problem in which a set $\mathcal{J}_i(\varphi) \subseteq J$ of locations is chosen to maximize the firm's profits $\pi_i(\varphi)$ in (9). We can alternatively express this problem as

$$\max_{I_{ij}\in\{0,1\}_{j=1}^{J}} \pi_i(\varphi, I_{i1}, I_{i2}, ..., I_{iJ}) = \varphi^{\sigma-1}\left(\gamma\sum_{j=1}^{J} I_{ij}T_j\left(\tau_{ij}w_j\right)^{-\theta}\right)^{(\sigma-1)/\theta} B_i - w_i\sum_{j=1}^{J} I_{ij}f_{ij}, \tag{10}$$

[11] These derivations are analogous to those performed by Eaton and Kortum (2002) to solve for the aggregate price index in their model of trade in final goods.

[12] See, among other, Halpern et al. (2011), Goldberg et al. (2010), and Gopinath and Neiman (2013). The cost function in (8) and profit function below in (9) are actually isomorphic (up to a scalar) to those derived from an Armington model in which inputs are differentiated by country of origin and their elasticity of substitution is constant and given by $(1+\theta)/\theta$.

where the indicator variable I_{ij} takes a value of 1 when $j \in \mathcal{J}_i(\varphi)$ and 0 otherwise. In theory, with knowledge of the small number of parameters that appear in (10), this problem could be solved computationally by calculating firm profits for different combinations of locations and picking the unique strategy yielding that highest level of profits. Nevertheless, this would amount to computing profits for 2^J possible strategies, which is clearly infeasible unless one chooses a small enough set J of candidate countries. In section 4, we will discuss alternative approaches to circumvent this dimensionality problem when estimating some of the parameters of the model. For the time being, we will focus on providing a characterization of some key analytic features of the solution to this optimal sourcing strategy problem.

The problem in (10) is not straightforward to solve because the decision to include a country j in the set $\mathcal{J}_i(\varphi)$ depends on the number and characteristics of the other countries in this set. That dependence is in turn crucially shaped by whether $(\sigma-1)/\theta$ is higher or lower than one. When $(\sigma-1)/\theta > 1$, the profit function $\pi_i(\varphi)$ features increasing differences in (I_{ij}, I_{ik}) for $j, k \in \{1, ..., J\}$ and $j \neq k$, and as a result the marginal gain from adding a new location to $\mathcal{J}_i(\varphi)$ is increasing in the number of elements in that set. This case is more likely to apply whenever demand is elastic and thus profits are particular responsive to variable cost reductions (high σ), and whenever input efficiency levels are relatively heterogeneous across markets (low θ), so that the expected reduction in costs achieved by adding an extra country into the set of active locations is relatively high. In the converse case in which $(\sigma-1)/\theta < 1$, we instead have that firm profits feature decreasing differences in (I_{ij}, I_{ik}), and the marginal gain from adding a new location to $\mathcal{J}_i(\varphi)$ is decreasing in the number of elements in that set. We shall refer to the case $(\sigma-1)/\theta > 1$ as the *complements* case, and to the case $(\sigma-1)/\theta < 1$ as the *substitutes* case.

Regardless of parameter values, we can establish that:

Proposition 1. The solution $I_{ij}(\varphi) \in \{0,1\}_{j=1}^{J}$ to the optimal sourcing problem (10) is such that a firm's sourcing capability $\Theta_i(\varphi) = \sum_{j=1}^{J} I_{ij}(\varphi) T_j (\tau_{ij} w_j)^{-\theta}$ is nondecreasing in φ.

Although this is not the focus of our analysis, it is worth pointing out that Proposition 1 has interesting implications for the size distribution of firms as implied by our model. Note that firm sales are given by a multiple σ of operating profits or

$$R_i(\varphi) = \sigma \varphi^{\sigma-1} (\gamma \Theta_i(\varphi))^{(\sigma-1)/\theta} B_i. \tag{11}$$

If the sourcing capability of firms were independent of φ, then the size distribution of firms would be governed by the distribution of the term $\varphi^{\sigma-1}$. Nevertheless, with $\Theta_i(\varphi)$ being nondecreasing in φ, the equilibrium size distribution of firms will feature more positive skewness than as implied by the distribution of $\varphi^{\sigma-1}$. For instance, if $G_i(\varphi)$ is a log-normal distribution with mean μ and variance σ^2, $\varphi^{\sigma-1}$ is also distributed log-normal with mean $(\sigma-1)\mu$ and variance $(\sigma-1)^2\sigma^2$. Yet, Proposition 1 suggests that the actual firm size distribution will exhibit a fatter right tail than as predicted by a log-normal distribution estimated on the whole population of firms. As a result, the

size distribution of the relatively large firms may be better approximated by distributions featuring higher positive skewness than the log-normal, such as the Pareto distribution. These implications of the model are broadly consistent with available empirical evidence on the size distribution of firms (see, for instance, Rossi-Hansberg and Wright, 2007, particularly their Figure 1).

It is important to emphasize that the result in Proposition 1 that a firm's sourcing capability is increasing in productivity does *not* imply that the extensive margin of sourcing at the firm level (i.e., the number of elements of $\mathcal{J}_i(\varphi)$) is necessarily increasing in firm productivity as well. For example, a highly productive firm from i might pay a large fixed cost to be able to offshore to a country j with a particularly high value of $T_j(\tau_{ij}w_j)^{-\theta}$ – thus greatly increasing Θ_i – after which the marginal incentive to add further locations might be greatly diminished in the substitutes case, that is whenever $(\sigma-1)/\theta < 1$.

Our description in section 2 of the stylized facts observed for U.S. firms suggests, however, that the extensive margin of offshoring *does* appear to respond positively to productivity. Furthermore, in Table 6 we provided strong evidence of the existence of a hierarchical structure in the sourcing location decisions of U.S. firms. In terms of our model, this would amount to the sourcing strategy $\mathcal{J}_i(\varphi)$ of relatively unproductive firms being a strict subset of the sourcing strategy of relatively productive firms, or $\mathcal{J}_i(\varphi_L) \subseteq \mathcal{J}_i(\varphi_H)$ for $\varphi_H \geq \varphi_L$. As we show in the Appendix, this strong pattern is in fact implied by our model in the complements case, and thus we have that:

Proposition 2. Whenever $(\sigma-1)/\theta > 1$, the solution $I_{ij}(\varphi) \in \{0,1\}_{j=1}^{J}$ to the optimal sourcing problem (10) is such that $\mathcal{J}_i(\varphi_L) \subseteq \mathcal{J}_i(\varphi_H)$ for $\varphi_H \geq \varphi_L$, where $\mathcal{J}_i(\varphi) = \{j : I_{ij}(\varphi) = 1\}$.

In the complements case, the model thus delivers a 'pecking order' in the extensive margin of offshoring which is reminiscent of the one typically obtained in models of exporting with heterogeneous firms, such as Eaton et al. (2011). Another implication of Proposition 2 is that, for firms with a sufficiently low value of core productivity φ, the set $\mathcal{J}_i(\varphi)$ may be a singleton, and the associated unique profitable location j of input production will necessarily be the one associated with the highest ratio $T_j(\tau_{ij}w_j)^{-\theta}/f_{ij}$. This in turn implies that if fixed costs of offshoring are disproportionately large relative to fixed costs of domestic sourcing, so that $f_{ii} \ll f_{ij}$ for any $j \neq i$, the model is consistent with the evidence in Table 6 of selection into offshoring – that is, the superior performance of firms engaged in offshoring relative to firms sourcing exclusively in their domestic economies.

The characterization results in Proposition 1 and 2 have proved useful for interpreting some of the stylized facts regarding the extensive margin of offshoring uncovered in section 2. It also follows from Proposition 2, that if one were to be able to rank countries by an index of their 'sourcing appeal', the measure of firms from i sourcing to a given country would be increasing in that index, and the computationally intensive problem (10) would be greatly simplified. Although it is obvious from (10) that such an index of sourcing appeal for a country j should be increasing in $T_j(\tau_{ij}w_j)^{-\theta}$ and decreasing in f_{ij}, it is less clear how *exactly* should such an index aggregate these variables.

We can obtain sharper characterizations of the solution to the sourcing strategy problem in (10) by making additional assumptions. Consider first a situation in which the fixed costs of offshoring

are common for all foreign countries, so $f_{ij} = f_{iO}$ for all $j \neq i$. In such a case, and regardless of the value of $(\sigma - 1)/\theta$, one could then rank foreign locations $j \neq i$ according to their sourcing potential $T_j (\tau_{ij} w_j)^{-\theta}$ and denote by $i_r = \{i_1, i_{2,}..., i_{J-1}\}$ the country with the r-th highest value of $T_j (\tau_{ij} w_j)^{-\theta}$. Having constructed i_r, it then follows that for any firm with productivity φ from i that offshores to at least one country, $i_1 \in \mathcal{J}_i(\varphi)$; for any firm that offshores to at least two countries, we have $i_2 \in \mathcal{J}_i(\varphi)$, and so on. In other words, not only does the extensive margin increase monotonically with firm productivity, but it does so in a manner uniquely determined by the ranking of the $T_j (\tau_{ij} w_j)^{-\theta}$ sourcing potential terms. It is important to emphasize that this result holds both in the complements case as well as in the substitutes case, though again it relies on the restrictive assumption of identical offshoring fixed costs across sourcing countries.

Even in the presence of cross-country differences in the fixed costs of offshoring, a similar sharp result emerges in the knife-edge case in which $(\sigma - 1)/\theta = 1$. In that case, the addition of an element to the set $\mathcal{J}_i(\varphi)$ has no effect on the decision to add any other element to the set, and the same pecking order pattern described in the previous paragraph applies, but when one ranks foreign locations according to the ratio $T_j (\tau_{ij} w_j)^{-\theta} / f_{ij}$ rather than $T_j (\tau_{ij} w_j)^{-\theta}$. This results is analogous to the one obtained in standard models of selection into exporting featuring constant marginal costs, in which the decision to service a given market is independent of that same decision in other markets.

We can also use the properties of the profit function to derive a few firm level comparative statics that hold constant the residual demand level B_i. First, and quite naturally, a reduction in any iceberg trade cost τ_{ij} or fixed cost of sourcing f_{ij} (weakly) increases the firm's sourcing capability $\Theta_i(\varphi)$ and thus firm-level profits. Second, in the complements case, a reduction of any τ_{ij} or f_{ij} also (weakly) increase the extensive margin of global sourcing, in the sense that the set $\mathcal{J}_i(\varphi)$ is nondecreasing in τ_{ij} and f_{ij} for any j. Finally, and perhaps more surprisingly, in the complements case it is also the case that a reduction of *any* τ_{ij} or f_{ij} (weakly) increase firm-level bilateral input purchases from *all* countries. Intuitively, in such a case, complementarities are strong enough to dominate the direct substitution effect related to market shares shifting towards the locations whose costs of sourcing have been reduced. It should be emphasized, however, that these results apply when holding the level of B_i fixed, when in fact it would be affected by these same parameter changes.

3.5 Aggregation and General Equilibrium

After having solved the sourcing strategy problem in (10), it is straightforward to derive the implications of the model for the aggregate volume of bilateral trade across countries. Because we have assumed that final goods are nontradable, we can focus on characterizing aggregate intermediate input trade flows between any two countries i and j. Given that firm spending on intermediate inputs constitutes a share $(\sigma - 1)/\sigma$ of revenue for all firms, we can use equation (11) and aggregate

across firms, to obtain aggregate imports from country j by firms based in i:

$$M_{ij} = (\sigma - 1) N_i B_i \int_{\tilde{\varphi}_{ij}}^{\infty} \chi_{ij}(\varphi) (\gamma \Theta_i(\varphi))^{(\sigma-1)/\theta} \varphi^{\sigma-1} dG_i(\varphi) . \quad (12)$$

In this expression, $\tilde{\varphi}_{ij}$ denotes the productivity of the least productive firm from i offshoring to j, while $\chi_{ij}(\varphi)$ is given in (6) for $j \in \mathcal{J}_i(\varphi)$ and by $\chi_{ij}(\varphi) = 0$ otherwise.

Consider now the general equilibrium of the model in which income equals spending and factor markets clear. We will first consider the case in which the set of firms in each country i is exogenously given by N_i, and in which the only sector hiring workers is manufacturing. With these assumptions, we can express the labor-market-clearing condition in country i as follows:

$$w_i L_i = (\sigma - 1) \sum_{j=1}^{J} N_j B_j \int_{\tilde{\varphi}_{ji}}^{\infty} \chi_{ji}(\varphi) (\gamma \Theta_j(\varphi))^{(\sigma-1)/\theta} \varphi^{\sigma-1} dG_j(\varphi) + N_i w_i \int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty} \sum_{k \in \mathcal{J}_i(\varphi)} f_{ik} dG_i(\varphi) . \quad (13)$$

The left-hand-side of equation (13) is the overall wage bill paid to workers in country i. This needs to equal to the sum of the remuneration paid to (production) workers from i employed in the production of intermediates by producers worldwide, and the remuneration to (non-production) workers from i employed by domestic firms when covering their fixed costs. Note that in the first term we have imposed that intermediates are produced by a competitive fringe of suppliers and thus the wage bill accruing to country i workers in the production of intermediates for firms worldwide is simply $\sum_j M_{ji}$ (see equation (12)). In the second term, $\vartheta(i)$ denotes the location from which the least productive active firm in country i sources its inputs, or formally, $\vartheta(i) = \{j \in J : \tilde{\varphi}_{ij} \leq \tilde{\varphi}_{ik}$ for all $k \in J\}$. As argued above, our evidence in section 2 suggests that typically $\vartheta(i) = i$.

The equality of income and spending, together with equations (3), (8) and constant-mark-up pricing, in turn imposes

$$B_i = \frac{1}{\sigma} \frac{w_i L_i + \Pi_i}{N_i \int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty} (\gamma \Theta_i(\varphi))^{(\sigma-1)/\theta} \varphi^{\sigma-1} dG_i(\varphi)}, \quad (14)$$

where aggregate profits are given by (see eq. (9))

$$\Pi_i = N_i \int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty} \left[\varphi^{\sigma-1} (\gamma \Theta_i(\varphi))^{(\sigma-1)/\theta} B_i - w_i \sum_{j \in \mathcal{J}_i(\varphi)} f_{ij} \right] dG_i(\varphi) . \quad (15)$$

So far we have discussed the general equilibrium of the model with a fixed number of firms, but we can easily extend the analysis to the case of endogenous entry. In such a case, equation (14)

continues to apply with $\Pi_i = 0$, while equation (15) is replaced by the free entry condition

$$\int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty} \left[\varphi^{\sigma-1} \left(\gamma \Theta_i \left(\varphi\right)\right)^{(\sigma-1)/\theta} B_i - w_i \sum_{j \in \mathcal{J}_i(\varphi)} f_{ij} \right] dG_i \left(\varphi\right) = w_i f_{ei}.$$

Combining these modified versions of equations (14) and (15), we find

$$B_i = w_i \frac{\int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty} \sum\limits_{j \in \mathcal{J}_i(\varphi)} f_{ij} dG_i \left(\varphi\right) + f_{ei}}{\int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty} \varphi^{\sigma-1} \left(\gamma \Theta_i \left(\varphi\right)\right)^{(\sigma-1)/\theta} dG_i \left(\varphi\right)} \tag{16}$$

and

$$N_i = \frac{L_i}{\sigma \left(\int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty} \sum\limits_{j \in \mathcal{J}_i(\varphi)} f_{ij} dG_i \left(\varphi\right) + f_{ei} \right)}. \tag{17}$$

Manipulating equations (13), (16) and (17), together with the definition of $\chi_{ji}(\varphi)$ in (6), we can then reduce wage determination to a system of non-linear equations characterized by

$$\sum_{j=1}^{J} \left[\frac{\int_{\tilde{\varphi}_{ji}}^{\infty} I_{ji}\left(\varphi\right) T_i \left(\tau_{ji} w_i\right)^{-\theta} \Theta_j \left(\varphi\right)^{(\sigma-1-\theta)/\theta} \varphi^{\sigma-1} dG_j \left(\varphi\right)}{\int_{\tilde{\varphi}_{j\vartheta(j)}}^{\infty} \Theta_j \left(\varphi\right)^{(\sigma-1)/\theta} \varphi^{\sigma-1} dG_j \left(\varphi\right)} \right] w_j L_j = w_i L_i \text{ for all } i \in J, \tag{18}$$

where remember that $\Theta_j(\varphi) \equiv \sum_{k \in \mathcal{J}_j(\varphi)} T_k (\tau_{jk} w_k)^{-\theta}$. In this expression, $I_{ji}(\varphi)$ is an indicator function taking a value of one when a firm from j with productivity φ has country i in its sourcing strategy (or $i \in \mathcal{J}_j(\varphi)$). Our results above imply that, in the complements case, $I_{ji}(\varphi) = 1$ for all $\varphi > \tilde{\varphi}_{ji}$, but the same is not necessarily true in the substitutes case.

Provided that equations (16)-(18) admit a unique solution for the triplet (w_i, N_i, B_i) for each country $i \in J$, it should be clear that the equilibrium will exist and will be unique. In particular, the firm-level combinatorial problem in (10) delivers a unique solution given w_i, B_i and exogenous parameters. This in turn determines all thresholds $\tilde{\varphi}_{ij}$ for any pair of countries (i, j) uniquely, and this in turn guarantees the existence of unique equilibrium measure of firms N_i.

The system of equations determining wages in (18) is highly nonlinear and a general proof of existence and uniqueness has eluded us to date. In the Appendix we show, however, that:

Proposition 3. Provided that the system in (18) admits a unique solution with positive wages in each country, equations (16) and (17) will in turn deliver a unique solution for B_i and N_i in each country $i \in J$.

The proof of this result is not entirely straightforward because – via the terms $\mathcal{J}_i(\varphi)$ and $\Theta_i(\varphi)$ – equations (16) and (17) are functions of *all* thresholds $\tilde{\varphi}_{ij}$ for all $(i, j) \in J \times J$, which in turn depend on the term B_i in ways that one is not able to characterize for general parameter values (see our discussion of the optimal sourcing strategy above). In the Appendix, we appeal to monotone comparative statics arguments to prove the existence and uniqueness of the equilibrium for a *given*

vector of equilibrium wages. It is worth emphasizing that this existence and uniqueness result holds both in the complements case as well as in the substitutes case.[13]

This *conditional* proof of existence and uniqueness is particularly useful when considering an alternative way to close the model that incorporates additional sectors of the economy. In particular, imagine that these other sectors jointly absorb a constant share $1 - \eta$ of the economy's income, and compete for labor with manufacturing firms. If $1 - \eta$ is large enough, the wage rate w_i in each country i will be pinned down in those sectors and can be treated as exogenous in solving for the equilibrium in the manufacturing sector. This would in turn imply that we can safely ignore the factor market clearing condition (18). Furthermore, if profits in these additional sectors are negligible, the only other necessary adjustment to the above equilibrium equations (16) and (17) is to multiply the numerator of (17) by η. It then follows from Proposition 3 that the general equilibrium of this multi-sector model exists and is unique.

3.6 Gravity

We begin this section with a special case of our framework that permits a comparison of the implications of our model for the structure of trade flows and welfare relative to those of some canonical models of trade with productivity heterogeneity. In particular, we shall begin by considering the case in which the fixed costs of offshoring are low enough to ensure that *all* firms acquire the capability to source inputs from *all* countries. This is obviously counterfactual in light of our stylized facts in section 2, but studying this unrealistic benchmark environment will prove to be useful below. To simplify matters, we will abstract from fixed costs of sourcing altogether and set them to zero.

When all firms import from everywhere, the optimal sourcing strategy of all firms in i is simply given by $J_i(\varphi) = \{1, 2, ..., J\}$, and both $\Theta_i = \sum_{k \in J} T_k (\tau_{jk} w_k)^{-\theta}$ and thus χ_{ij} are independent of φ. This allows to greatly simplify the equilibrium equations (16), (17), and (18). The latter equation, in particular, reduces to the much simpler system

$$\sum_{j=1}^{J} \frac{T_i (\tau_{ji} w_i)^{-\theta}}{\Theta_j(\varphi)} w_j L_j = w_i L_i \text{ for all } i \in J. \tag{19}$$

Furthermore, the volume of imports (of intermediates) from country j by firms from i in equation (12) can be written, appealing to (16) and (17), as

$$M_{ij} = \frac{(\sigma - 1)}{\sigma} \frac{T_j (\tau_{ij} w_j)^{-\theta}}{\Theta_i} w_i L_i. \tag{20}$$

We can re-express equation (20) in a more familiar form by aggregating j's exports over importing

[13] Although we omit the proof, it is straightforward to show, following analogous steps to those in the Appendix, that in the case of a fixed number of firms N_i, the equilibrium given a vector of wages also exists and is unique.

countries i (i.e., computing $\sum_i M_{ij}$) and using (19) to simplify to

$$M_{ij} = \frac{(\sigma-1)}{\sigma}\tau_{ij}^{-\theta}\frac{w_i L_i}{\Theta_i}\frac{w_j L_j}{\sum_k \tau_{kj}^{-\theta}\frac{w_k L_k}{\Theta_k}}. \tag{21}$$

This is a standard gravity equation relating bilateral trade flows to bilateral trade barriers τ_{ij}, the ratio between the importer country's GDP $w_i L_i$ and this country's sourcing capability Θ_i (which is a negative transform of that country's ideal price index), and the exporter country's GDP $(w_j L_j)$ divided by a structural measure of this exporting country's sourcing appeal (the summation term in the denominator).[14] Notice that equation (21) structurally justifies the use of an empirical log-linear specification for bilateral trade flows with importer and exporter asymmetric fixed effects and measures of bilateral trade frictions τ_{ij}. Furthermore, the model indicates that the elasticity of trade flows with respect to changes in these bilateral trade frictions is shaped by the Fréchet parameter θ, just as in the Eaton and Kortum (2002) framework. The intuition for this result is analogous to our earlier discussion of the effects of bilateral trade frictions on the cross-section of import purchases at the firm level. This should not be surprising since, in the absence of selection into offshoring, all firms buy inputs from all markets according to the same market shares χ_{ij}.

How does the introduction of fixed costs of sourcing and selection into offshoring affect the gravity specification in (21)? Following similar steps as in the derivation of (21), we find that the volume of imports from country j by firms from i is now given by

$$M_{ij} = \frac{(\sigma-1)}{\sigma}\tau_{ij}^{-\theta}\Lambda_{ij}\frac{w_i L_i}{P_i^{1-\sigma}/N_i}\frac{w_j L_j}{\sum_k \tau_{kj}^{-\theta}\Lambda_{kj}\frac{w_k L_k}{P_j^{1-\sigma}/N_j}} \tag{22}$$

where

$$\Lambda_{ij} = \int_{\tilde{\varphi}_{ij}}^{\infty} I_{ij}(\varphi)\left(\Theta_i(\varphi)\right)^{(\sigma-1-\theta)/\theta}\varphi^{\sigma-1}dG_i(\varphi), \tag{23}$$

and where, using the cost function in (8) and constant markup pricing, country i's price index can be expressed as

$$P_i^{1-\sigma} = \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} N_i \int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty}\left(\gamma\Theta_i(\varphi)\right)^{(\sigma-1)/\theta}\varphi^{\sigma-1}dG_i(\varphi). \tag{24}$$

Equation (22) is identical to the gravity equation in (21) except for the fact that $N_i\Lambda_{ij}/P_i^{1-\sigma}$ and $N_k\Lambda_{kj}/P_k^{1-\sigma}$ replace Θ_i and Θ_k, respectively. It is also apparent from (23) that when Θ_i is independent of φ and there is no selection into offshoring (so $\tilde{\varphi}_{ij} = \tilde{\varphi}_{i\vartheta(i)}$), we have $N_i\Lambda_{ij}/P_i^{1-\sigma} = \left(\frac{\sigma}{\sigma-1}\right)^{\sigma-1}\Theta_i$ for all j and the two equations become identical.

As similar as these two equations might appear, they carry significantly different implications. Notice, in particular, that even after partialling out importer and exporter fixed effects, we are

[14] The price index in country i is given by $P_i^{1-\sigma} = N_i\int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty} p_i(\varphi)^{1-\sigma}dG_i(\varphi)$. In the absence of selection into importing, this reduces to $P_j = \left(\frac{\sigma}{\sigma-1}\right)N_i(\gamma\Theta_i)^{-1/\theta}$, where N_i is given in (17).

left with the fact that both τ_{ij} as well as Λ_{ij} in (23) vary across exporters even when fixing the importer. Furthermore, the lower bound $\tilde{\varphi}_{ij}$ in the integral of Λ_{ij} will naturally depend on bilateral trade costs τ_{ij} between i and j and thus the elasticity of trade flows with respect to trade frictions is no longer simply given by $-\theta$. In particular, if as one might expect, $\tilde{\varphi}_{ij}$ is increasing in τ_{ij}, reductions in trade costs will generate a magnified positive effect on trade volumes via increased selection into offshoring. This second effect is again an extensive margin effect but at the *firm-level* rather than the product-level discussed above and present also in Eaton and Kortum (2002).

A second key difference between the two specifications is that as long as $(\sigma-1)/\theta \neq 1$, Λ_{ij} in (23) will be a function of $\Theta_i(\varphi)$ for $\varphi > \tilde{\varphi}_{ij}$, and will thus depend on technology, trade costs and wages in all countries, and not just i and j. Hence, equation (22) is an *extended* gravity equation – to use the term in Morales et al. (2014) – featuring third market effects. The sign of these effects in turn appears to crucially depend on whether we are in the complements or substitutes case.

It is interesting to briefly consider the knife-edge case $(\sigma-1)/\theta = 1$ because it shuts down these third market effects, while keeping active the country-level extensive margin of offshoring. It will also prove convenient, as is typical in models with monopolistic competition and heterogeneous firms, to assume that firms draw their core productivity from a Pareto distribution with shape parameter $\kappa > \sigma - 1$, or $G_i(\varphi) = 1 - (\underline{\varphi}_i/\varphi)^{\kappa}$. In such a case, and after cumbersome manipulations (see the Appendix) we find that equation (22) simplifies to

$$M_{ij} = \frac{(\sigma-1)}{\sigma} \tau_{ij}^{-\kappa} f_{ij}^{1-\kappa/(\sigma-1)} \Psi_i \frac{w_i L_i}{P_i^{-\kappa}} \frac{w_j L_j}{\sum_k \tau_{kj}^{-\kappa} f_{kj}^{1-\kappa/(\sigma-1)} \Psi_k \frac{w_k L_k}{P_k^{-\kappa}}}, \tag{25}$$

where

$$\Psi_i = \underline{\varphi}_i^{\kappa} L_i^{\kappa/(\sigma-1)} / f_{ei}.$$

is a term that will be absorbed by an importer fixed effect.[15] As anticipated, this equation is indeed a gravity equation (without third market effects) but with a trade elasticity $\kappa > \sigma - 1 = \theta$ higher than the one obtaining when the model features no country-level extensive margin.

It may be surprising that the Fréchet parameter θ, which was key in governing the 'trade elasticity' (i.e., the elasticity of trade flows to variable trade costs) at the firm level, is now irrelevant when computing that same elasticity at the aggregate level. To understand this result it is useful to relate our framework to the multi-country versions of the Melitz model in Chaney (2008), Arkolakis et al. (2008) or Helpman et al. (2008). In those models, firms pay fixed costs of exporting to obtain additional operating profit flows proportional to $\varphi^{\sigma-1}$ that enter linearly and separably in the firm's profit function. Even though in our model, selection into offshoring increases firm profits by reducing effective marginal costs, whenever $\sigma - 1 = \theta$, the gain from adding a new market is strictly

[15] In this derivation, we have used the fact that whenever $(\sigma-1)/\theta = 1$, the extensive margin sourcing decision can be studied independently for each country, and we thus have that $\tilde{\varphi}_{ij}$ must satisfy

$$\tilde{\varphi}_{ij}^{\sigma-1} \left(\gamma T_j (w_j \tau_{ij})^{-\theta} \right) B_i = w_i f_{ij}.$$

separable in the profit function and also proportional to $\varphi^{\sigma-1}$. Hence, this effect is isomorphic to the firm obtaining additional sale revenue by selecting into exporting. It is thus not surprising that the gravity equation we obtain in (25) is essentially identical to those obtained by Chaney (2008) or Arkolakis et al. (2008).

It is important to bear in mind, however, that $\sigma - 1 = \theta$ is a knife-edge case. Our results above demonstrate that our model generally delivers a trade elasticity that is not uniquely pinned down by exogenous parameters as well as an extended gravity equation with third market effects.

3.7 Welfare

We next explore the welfare implications of trade opening in our framework and how they compare to those in other related models. As we did in the last section, it is useful to consider first the benchmark case in which all fixed costs of sourcing are set to 0 and the model features no country-level extensive margin of sourcing. As we noted above, our model in that case shares many features with the Eaton and Kortum (2002) framework and thus it should not come as a surprise that we can obtain a formula for welfare similar to that in their paper.

More specifically, we can use the price index equation (24) together with $\chi_{ii} = T_i (w_i)^{-\theta} / \Theta_i$ and equation (17), to express real income, and hence welfare, in country i as

$$\frac{w_i}{P_i} = \frac{\sigma - 1}{\sigma} \left(\frac{\gamma T_i}{\chi_{ii}} \right)^{1/\theta} \left(\frac{L_i}{\sigma f_e} \right)^{1/(\sigma-1)} \left(\int_{\underline{\varphi}_i}^{\infty} \varphi^{\sigma-1} dG_i(\varphi) \right)^{1/(\sigma-1)}. \tag{26}$$

It is straightforward to verify that as $\sigma \to \infty$, and the final-good sector becomes competitive, this expression converges to the one in Eaton and Kortum (2002), and the share of spending on domestic inputs and the elasticity of trade to variable trade costs are sufficient statistics for welfare (see also Arkolakis et al., 2012). For bounded σ, our model features an additional scale effect not present in the Eaton and Kortum (2002) framework – the term $L_i^{1/(\sigma-1)}$ –, which is associated with love-for-variety in final-good demand, as in Krugman (1980) or Melitz (2003). In fact, when $\theta \to \infty$ and thus we eliminate the variability in the Fréchet draws, welfare is identical to that in the closed-economy version of Melitz (2003), or to that in a version of Krugman (1980) in which firms are identical and have a common productivity level equal to $\hat{\varphi} = \left(\int_{\underline{\varphi}_i}^{\infty} \varphi^{\sigma-1} dG_i(\varphi) \right)^{1/(\sigma-1)}$.

Unlike in Krugman (1980) and Melitz (2003), this scale effect will be unaffected by trade opening because final goods are assumed to be nontradable. As a result, the only effect of trade on welfare is to reduce the domestic trade share χ_{ii}, just as in Eaton and Kortum (2002). Furthermore, because $\chi_{ii} = 1$ under autarky, the welfare gains from trade relative to (counterfactual) autarky can easily be read off the formula with knowledge of the value of χ_{ii} with trade and of θ, a point emphasized by Arkolakis et al. (2012). The gains from trade are naturally decreasing in θ, because the variance of input productivities is decreasing in θ, hence making comparative advantage more powerful when θ is low.

So far, we have focused on the special case in which all fixed costs of sourcing are set to zero.

How does welfare formula change when selection into offshoring becomes operative? In such a case, and assuming a Pareto distribution of productivity for final good producers, we find that

$$\frac{w_i}{P_i} = \frac{\sigma-1}{\sigma}\left(\frac{\gamma T_i}{\chi_{ii}^{agg}}\right)^{1/\theta}\left(\frac{(\sigma-1)L_i}{\sigma\kappa f_e}\right)^{1/(\sigma-1)} \times \left(\frac{\left(\int_{\tilde{\varphi}_{ii}}^{\infty} I_{ii}(\varphi)\,\Theta_i(\varphi)^{(\sigma-1)/\theta-1}\varphi^{\sigma-1}dG_i(\varphi)\right)^{(\sigma-1)/\theta}}{\left(\int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty}(\Theta_i(\varphi))^{(\sigma-1)/\theta}\varphi^{\sigma-1}dG_i(\varphi)\right)^{(\sigma-1)/\theta-1}}\right)^{1/(\sigma-1)} \tag{27}$$

where χ_{ii}^{agg} is the *aggregate* share of spending on domestic inputs (note that $\chi_{ii}(\varphi)$ varies across firms now), and is given by

$$\chi_{ii}^{agg} = \frac{T_i(w_i)^{-\theta}\int_{\tilde{\varphi}_{ii}}^{\infty} I_{ii}(\varphi)\,\Theta_i(\varphi)^{(\sigma-1)/\theta-1}\varphi^{\sigma-1}dG_i(\varphi)}{\int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty}\Theta_i(\varphi)^{(\sigma-1)/\theta}\varphi^{\sigma-1}dG_i(\varphi)}.$$

Equation (27) is similar to (26) but it incorporates a new terms that is endogenously affected by trade opening. Hence, the own trade share χ_{ii}^{agg} and the firm-level trade elasticity θ are no longer sufficient statistics for the welfare effects of trade. This should not be particularly surprising given that we have shown above that the aggregate elasticity of trade flows to variable trade frictions is no longer given by θ when the model features and extensive margin of offshoring.

Indeed, when we consider the knife-edge case in which $\sigma-1=\theta$, we have shown above that the aggregate trade elasticity becomes $\kappa>\theta$, and the formula for aggregate welfare can be reduced after several calculations (see the Appendix), to

$$\frac{w_i}{P_i} = \frac{\sigma-1}{\sigma}(\gamma T_i L_i)^{1/(\sigma-1)}\left(\chi_{ii}^{agg}\right)^{-1/\kappa}\left(\frac{\underline{\varphi}_i^{\kappa} f_{ii}^{1-\kappa/(\sigma-1)}}{\sigma^{\kappa/(\sigma-1)}f_{ei}}\frac{\sigma-1}{\kappa-\sigma+1}\right)^{1/\kappa}. \tag{28}$$

This formula is identical to the welfare expression in Arkolakis et al. (2008) except for the presence of the term $(\gamma T_i)^{1/(\sigma-1)}$. The only term in (28) that is endogenously affected by changes in trade costs is $(\chi_{ii}^{agg})^{-1/\kappa}$ and thus we recover the result in Arkolakis et al. (2008) or Arkolakis et al. (2012) than the own trade share χ_{ii}^{agg} and the aggregate trade elasticity κ are sufficient statistics for welfare.

As shown in (27), however, when departing from the knife-edge case $\sigma-1=\theta$, evaluating the welfare gains from trade is much more complicated and requires computing the equilibrium of the model, an approach we follow in later sections.

3.8 An Extension: Tradable Final Goods

We have assumed so far that final-good varieties are prohibitively costly to trade across borders. We have done so to focus our analysis on the determinants and implications of selection into global sourcing. In this section, we briefly relax this assumption and demonstrate the existence of intuitive complementarities between the extensive margin of exporting and that of importing at the firm level.

Suppose then that trade in final-varieties is only partially costly and involves both iceberg trade costs τ_{ij}^X as well as fixed costs f_{ij}^X of exporting. Firm behavior conditional on a sourcing strategy is largely analogous to that in section 3.3. In particular, after observing the realization of the supplier productivity shocks, each final-good producer will continue to choose the location of production for each input production to minimize costs, which will lead to the same marginal cost function $c_i(\varphi)$ obtained above in equation (8). The main novelty is that the firm will now produce output not only for the domestic market but also for a set of endogenously chosen foreign markets, which constitute the firm's 'exporting strategy'. We can then express the problem of determining the optimal exporting and sourcing strategies of a firm from country i with core productivity φ as:

$$\begin{aligned}\max_{\substack{I_{ij}^M \in \{0,1\}_{j=1}^J \\ I_{ik}^X \in \{0,1\}_{k=1}^J}} \pi_i\left(\varphi, \mathbf{I}^M, \mathbf{I}^X\right) &= \varphi^{(\sigma-1)}\left(\gamma \sum_{j=1}^J I_{ij}^M T_j \left(\tau_{ij} w_j\right)^{-\theta}\right)^{(\sigma-1)/\theta} \sum_{k=1}^J I_{ik}^X \left(\tau_{ik}^X\right)^{1-\sigma} B_k \\ &\quad - w_i \sum_{j=1}^J I_{ij}^M f_{ij} - w_i \sum_{k=1}^J I_{ik}^X f_{ij}^X,\end{aligned}$$

Note that $\mathbf{I}$ and X denote the vector of extensive margin import and export decisions, respectively. It is straightforward to see that, whenever $(\sigma-1)/\theta > 1$, this more general profit function continues to feature increasing differences in $\left(I_j^M, I_k^M\right)$ for $j,k \in \{1,...,J\}$ with $j \neq k$, and also features increasing differences in $\left(I_j^M, \varphi\right)$ for any $j \in \{1,...,J\}$. As a result, Proposition 2 continues to apply here and we obtain a 'pecking order' in the extensive margin of offshoring in the complements case.

The key new feature of the above profit function $\pi_i\left(\varphi, \mathbf{I}^M, \mathbf{I}^X\right)$ is that it also features increasing differences in (I_j, X_k) for any $j,k \in \{1,...,J\}$ and increasing differences in (X_j, φ) for any $j \in \{1,...,J\}$. This has at least two implications. First, regardless of whether $\sigma - 1 > \theta$ or $\sigma - 1 < \theta$, any change in parameters that increases the sourcing capability $\Theta_i(\varphi)$ of the firm – such as reduction in any τ_{ij} or an increase in any T_j – will necessarily lead to a (weak) increase in the vector $\mathbf{I}^X$, and thus (weakly) increase the export margin of exporting. Second, restricting attention to the complements case $(\sigma-1)/\theta > 1$, the model delivers a complementarity between the exporting and importing margins of firms. For instance, holding constant the vector of residual demand parameters B_i, reductions in the costs of trading final goods across countries will not only increase the participation of firms in export markets, but will also increase the extensive margin of sourcing, in the sense that vector $\mathbf{I}^M$ is non-increasing in τ_{ik}^X. Furthermore, as firm productivity increases, the participation of firms in both export and import markets increases, and at a faster rate than when the export margin is shut down. This result complementarity result is useful in interpreting the fact that, as Bernard et al. (2007) indicate, 41 percent of U.S. exporting firms also import while 79 percent of importers also export. Although it would be interesting to explore the aggregate implications of this extended framework, we shall not do so in this paper because our quantitative

exercises in the next section does not incorporate these features into the analysis.

4 Empirical Implementation

In the theory section, we provide a parsimonious model which characterizes the margins of firms' import sourcing decisions. The model is consistent with the patterns of U.S. firms' import behavior presented in section 2. In this section, we use the same firm-level data to structurally estimate the key parameters of our model. Following our model, each firm's decision can be separated into two steps: (1) Choosing its sourcing strategy and (2) Purchasing the optimal amount of intermediate inputs from each country contained in the sourcing strategy. We exploit the assumption of a continuum of intermediate inputs and full support in the product-country-specific productivity shocks to interpret a firm's chosen sourcing strategy as the set of countries from which the firm has positive imports.[16] Our theory implies a gravity equation for a firm's import purchases given its sourcing strategy (and zero import purchases from countries not contained in this set). In the first step of the empirical implementation, we use a simple linear regression to obtain an estimate of each country's sourcing potential $T_j\left(\tau_{ij}w_j\right)^{-\theta}$ from a U.S. perspective (i.e., $i = U.S.$). In the second step, we estimate the productivity dispersion parameter, θ, by projecting the estimated sourcing potential values on observed cost shifters and other controls, as well as measure the elasticity of demand, σ from observed variable mark-ups. We find robustly across various specifications that the ratio of the parameter estimates implies that the marginal benefit of importing from one country is increasing in the set of other countries in a firm's sourcing strategy. In the third and final step, we estimate the fixed cost of sourcing and other distributional parameters via simulated method of moments. We apply techniques developed by Jia (2008), originally designed to estimate an entry game among chains and other discount retailers in a large number of markets, to the sourcing strategy choices of U.S. firms.

Because we use data on the sourcing strategies of firms from a single country, in what follows, we often drop the subscript i from the notation, with the understanding that the unique importing country is the U.S. We also denote a firm by superscript n. The data we use for the structural estimation are based on the same set of manufacturing firms described in section 2.

We construct a measure of a firm's total intermediate input purchases based on the difference between the firm's sales and its value-added, adjusting for changes in its inventories of final goods and materials. This approach ensures a more complete metric of a firm's inputs than traditional measures based purely on manufacturers' use of materials because it takes into account the input usage of both the manufacturing as well as the wholesale establishments of U.S. firms.[17] The model

[16] Manufacturing firms tend to use a large number of tasks and products in the production of their goods. Therefore we think it is plausible that if a firm has paid the fixed cost of sourcing from a given country, there will be a least one input for which that country is the lowest cost supplier.

[17] The wholesale sector includes a significant number of plants that design goods and coordinate production, often by offshoring, but do not perform physical transformation activities (see Bernard and Fort, 2013, for a description). Ignoring these plants' inputs could severely understate multi-sector firms' total inputs. For example, Feenstra and Jensen (2012) find that a significant fraction of some manufacturing firms' imports are not reported as input purchases.

does not take a stance on whether intermediate inputs are sourced within or across firm boundaries. For the purposes of this paper, this is of little relevance for international transactions, but it might lead to important biases in our measure of overall input use if a significant share of domestic inputs is produced within the firm and is recorded as value added. For this reason, we add a firm's total production-worker wage bill to the difference between the firm's sales and value added. In terms of our model, this corresponds to assuming that the final-good producer employs production workers to manufacture internally any inputs produced by the firm, while it uses the other factors of production (nonproduction workers, physical capital, and land) to combine intermediate inputs and cover all fixed costs. This approach is also motivated by the notion that the services typically provided by production workers are particularly offshorable. This new measure of total intermediate input purchases is highly correlated with traditional input measures for manufacturing firms based on reported inputs of materials and parts. A firm's share of inputs from country j, χ_{ij}, is computed as imports from j divided by total input purchases. A firm's total use of domestic inputs, χ_{ii}, is simply the difference between its total input purchases and imports, divided by total input purchases.

To facilitate the estimation, we include only those countries that have at least 200 U.S. firms importing from them. This criterion leaves us with a total of 64 foreign sourcing options for firms. We include firms that import from these countries in the estimation, but adjust their input usage by subtracting their imports from the excluded countries.[18]

4.1 Step 1: Estimation of a Country's Sourcing Potential

Equation (6) in the model specifies the share of inputs a firm will source from country j, given its sourcing strategy $\mathcal{J}_i$. Country j's sourcing potential – from the perspective of country i – is summarized by the term $\xi_j = T_j (\tau_{ij} w_j)^{-\theta}$. We normalize the domestic sourcing potential term, $\xi_i = 1$.[19] Rearranging equation (6) by taking logs and subtracting the domestic share yields the following log-linear linear expression for any firm n

$$\log \chi_{ij}^n - \log \chi_{ii}^n = \log \xi_j + \log \epsilon_j^n. \tag{29}$$

In order to turn the model's equilibrium condition (6) into an empirical specification, note that this equation includes a firm-country-specific shock ϵ_j^n (relative to the firm-domestic-specific shock which we set at $\epsilon_i^n = 1$). Intuitively, this specification will allow us to identify a country's average sourcing potential ξ_j by observing how much a firm imports from that country relative to the same

We address this issue by including a firm's wholesale plants' inputs. Although there is no way to measure inputs for establishments outside the manufacturing and wholesale sectors, those plants are much less likely to be involved in production or importing. An alternative approach would be to use an estimate of the demand elasticity σ and exploit the CES structure of our model to back out input usage from sales data.

[18]We note that a very small fraction of firms has negative values for its domestic input purchases. This occurs when a firm's total input purchases are less than its imports. Likely explanations for this are measurement error and imports of capital equipment. We drop these firms from the estimation since we can use observations with negative input purchases. Since these firms are a small fraction of our total sample, we do not report how many there are to avoid future disclosure problems. We will provide details on this in the final version of the paper.

[19]We set domestic trade costs to 1, express all wages in terms relative to the U.S. wage, and fix $T_i = 1$.

firm's domestic input purchases, restricting attention to countries included in the firm's sourcing strategy. For this measurement strategy to be consistent, it is important that there be no selection based on the errors in this regression. This condition will be satisfied, for example, if firms only learn the country-specific efficiency shocks ϵ_j^n after their sourcing strategy is selected, or even more simply, if the term ϵ_j^n simply represents measurement error.[20]

Table 7 provides summary statistics from estimating equation 29 via OLS, using country fixed effects to capture the ξ_j terms. The estimated coefficients on these fixed effects represent each country's sourcing potential. All sourcing potential fixed effects are significant at the 99 percent level. We have also estimated these sourcing potential measures controlling for industry effects. The estimates are highly correlated (0.996) with our baseline results and retain their statistical significance. To simply the interpretation of the fixed effects, we estimate equation 29 without a constant. The adjusted R^2 from this specification is 0.85. The same specification with a constant yields an adjusted R^2 of 0.09. Including industry controls raises the R^2 to 0.15. We note that these relatively low R^2s are typical when using firm-level data with high levels of idiosyncractic noise.

Figure 2 plots the estimated sourcing potential fixed effects against total input purchases (left panel) and against the number of firms importing from that country (right panel). Our parameter estimates suggest that China has the highest sourcing potential for U.S. firms, followed by Canada and Taiwan. More firms import from Germany and United Kingdom than from Taiwan, however, and more firms import from Canada than from China, suggesting that fixed costs of sourcing are likely to differ across source countries. Despite some heterogeneity, the number of firms and total import purchases are clearly positively associated with a country's sourcing potential, with a tighter relationship between the sourcing potential and the number of firms sourcing from a country.

Table 7: Summary statistics for sourcing potential estimation

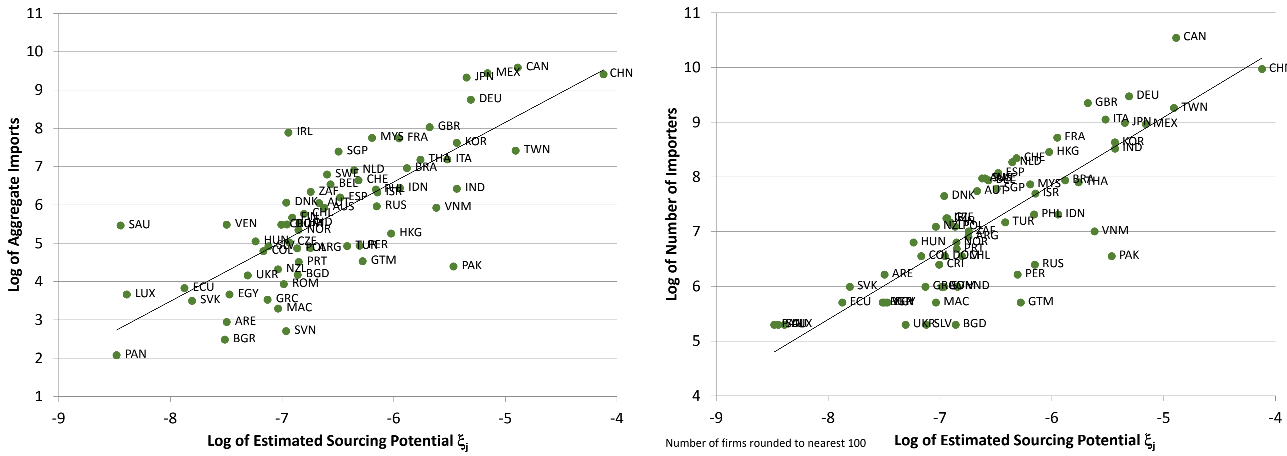
Number of observations	200,000
Number of importing firms	64,600
Mean Squared Error	2.64
Range of foreign $\log \xi_j$	- 4.12 to -8.42
Sum of foreign ξ_j	0.137

Notes: Summary statistics for regression based on equation (29). Estimated Fixed effects are displayed in Figure 2. Number of observations rounded for disclosure avoidance. The R^2 with a constant is 0.09.

Our estimates of the sourcing potential of a country enable us to calculate the extent to which the sourcing capability of a firm $\Theta^n = \sum_{j \in \mathcal{J}^n} \xi_j$ is higher if it imports from all countries as opposed to sourcing only domestically. Since the domestic sourcing potential was normalized to one and the summation of the foreign sourcing potential terms is 0.137, these results imply the sourcing

[20] However, we need to rule out measurement error related to a firm's global sourcing strategy. In other words, we assume throughout that the set of countries from which the firm imports is correctly observed by the researcher.

Figure 2: Country sourcing potential parameters



capability of a firm that sources from all 64 countries is 13.7 percent larger than that of a purely domestic sourcing firm. How much this higher sourcing capability lowers marginal cost depends according to equation (8) on the dispersion parameter θ of the intermediates productivities.

Remember that this parameter θ, together with the elasticity of demand σ faced by the U.S. importer, are crucial for determining the qualitative implications of our model for the optimal choice of a global sourcing strategy. We next turn to discussing how we estimate these two parameters.

4.2 Step 2: Estimation of the Elasticity of Demand and Input Productivity Dispersion

4.2.1 Estimation of Elasticity of Demand

It is simpler to start by discussing how we recover σ from the data. Note that with CES preferences and monopolistic competition, the ratio of sales to variable input purchases (including intermediates and basic factors of production) is $\sigma/(\sigma-1)$. We exploit this relationship to obtain a parameter value for σ by calculating a measure of average mark-ups from the establishment-level data in the Census of Manufactures. Specifically, the mark-up is the ratio of sales to variable inputs, where inputs are the sum of an establishment's materials, wages, capital expenditures, and total expenses. The mark-up for the median establishment is 35 percent, with a bootstrapped standard error of 0.88. This implies an estimate for the elasticity of demand, σ, of 3.85.

4.2.2 Estimation of Dispersion of Input Productivity Shocks

A second key parameter of our model is the dispersion of the productivity shocks of the intermediate inputs. Conditional on the firm's sourcing strategy, θ represents the firm-level trade elasticity in our model. We next use data on cost-shifters – wages and tariffs – to identify this elasticity. Recall that the sourcing potential ξ_j which we estimated in the previous section, is a function of a country's technology parameter, trade costs, and wages (apart from the parameter θ). We thus project the

estimated sourcing potential on proxies for all these terms, including R&D stock, capital per worker, a measure of control of corruption, wages, distance, tariffs, and common language.[21] Specifically, we estimate the following equation:

$$\begin{aligned}\log \xi_j = &\beta_r \log \text{R\&D}_j + \beta_k \log \text{capital}_j + \beta_C \text{control of corruption} - \theta \log w_j \\ &- \theta \left(\log \beta_c + \beta_d \log \text{distance}_{ij} + \text{language}_{ij} \log \beta_l + \log(1 + \text{tariff}_{ij})\right) + \iota_j. \end{aligned} \tag{30}$$

Notice that the parameter θ can be recovered from the coefficient associated with wages or tariffs in that regression. A potential issue with the use of country wage data is the fact that variation in wages partly reflects differences in worker productivity and skill across countries. Since firms' sourcing decisions are based on the cost of an efficiency unit of labor, we follow Eaton and Kortum (2002) and use a human capital-adjusted wage. We measure country human capital using the Hall and Jones (1999) methodology with updated data from Barro and Lee (2010). Even adjusting for skill differences across workers, there are other country-level factors that are likely correlated with the average wage, such as infrastructure, that will lead to an upward bias on the wage coefficient. To address this issue, as well the potential for measurement error, we instrument for a country's wage using its population (also from Barro and Lee (2010)).

The first column of Table 8 presents the OLS estimate of θ obtained when running (30) while constraining the same coefficient to apply to wages and tariffs. Column 2 provides the analogous IV estimate of θ using population as an instrument. Notice that the IV estimate (-1.785) is, as expected, larger in absolute value than the OLS estimate. It is interesting that, in line with our discussion in section 3.6, the data on firm-level trade flows suggest a much larger dispersion in productivities across countries than as typically obtained with aggregate trade data. For example, Eaton and Kortum (2002) estimate a coefficient of -3.8 using data on wages. Similarly to them, we find a coefficient of -4.763 when using the same specification as in equation (30) but with aggregate imports as a left-hand-side variable. These results are displayed in columns 4 and 5. It is noteworthy that our estimate of θ is sensitive to whether a measure of control of corruption is included in the projection or not. Corruption may be correlated with the extent to which incomplete contracting affects U.S. firms sourcing decisions (see Antràs (2014) for some evidence). The results without control of corruption lead to a lower estimate for θ of -1.083, as displayed in column 3. The aggregate trade elasticity in this specification (see column 6) is also lower with an estimate of -2.399. We think that controlling for corruption in this regression makes sense, but since this is not a commonly used control in the literature which estimates the trade elasticity, we highlight the difference if a specification without this control is used.

These estimates imply that a firm that sources from all countries faces between 7 percent ($1.137^{(-1/1.78)}$) and 11 percent ($1.137^{(-1/1.08)}$) lower input cost than a firm sourcing purely do-

[21] The country R&D data are from the World Bank Development Indicators. Physical capital is based on the methodology in Hall and Jones (1999), but constructed using the most recent data from the Penn World Tables. Control of corruption is from the World Bank's Worldwide Governance Indicators. The wage data are from the ILO data described by Oostendorp (2005). Tariffs are the simple average of country tariffs from the World Bank WITS database. Distance and language are from CEPII.

Table 8: Estimation of with-in firm and aggregate trade elasticity

	log ξ			log aggregate import		
	OLS	IV	IV	OLS	IV	IV
log(1+tariff) + log wage	-0.471	-1.785	-1.083	-0.516	-4.763	-2.399
	(0.175)	(0.651)	(0.320)	(0.389)	(1.824)	(0.788)
log distance	-0.326	-0.730	-0.514	-1.025	-2.329	-1.614
	(0.182)	(0.311)	(0.217)	(0.404)	(0.870)	(0.534)
common language	0.291	0.199	0.434	0.489	0.189	0.965
	(0.209)	(0.288)	(0.219)	(0.463)	(0.806)	(0.539)
log R&D	0.372	0.490	0.445	0.668	1.052	0.859
	(0.0500)	(0.0875)	(0.0556)	(0.111)	(0.245)	(0.137)
log KL	-0.158	0.506	0.415	-0.266	1.879	1.528
	(0.170)	(0.384)	(0.291)	(0.378)	(1.076)	(0.717)
Control of corruption	0.117	0.626		0.323	1.966	
	(0.147)	(0.308)		(0.327)	(0.864)	
Constant	-7.034	-11.44	-9.925	6.282	-7.969	-2.680
	(0.852)	(2.340)	(1.523)	(1.891)	(6.556)	(3.752)
Observations	57	57	58	57	57	58

Notes: Standard errors in parentheses. IVs are population and tariff.

mestically, and consequently its sales are between 22 percent ($1.137^{(-2.85/1.78)}$) and 40 percent ($1.137^{(-2.85/1.08)}$) larger.

Across various specifications that we have experimented with, one finds robustly that the ratio of elasticity of demand, $\sigma - 1$, to the dispersion of intermediate good efficiencies, θ, is larger than one. As argued in section 3, this implies that the profit function has increasing differences in the various sourcing decisions. As explained below, in the third step of our estimation, we exploit this feature of our model when numerically solving the firm's problem in order to estimate the fixed cost of sourcing associated with different markets. For the following, we use the θ estimate of 1.78.

4.3 Step 3: Estimation of Fixed Costs of Sourcing

We estimate fixed cost of sourcing and other distributional parameters via Simulated Method of Moments. In order to match the model outcomes to moments generated from firm-level data, we extend our model to incorporate, aside from core productivity differences, two additional sources of firm heterogeneity: (i) firm-country-specific variation in fixed cost of sourcing, and (ii) firm-country-specific shocks to the sourcing potential, ϵ_j^n (which we already introduced in the estimation of sourcing potentials above). We assume that the firm initially observes its set of fixed cost draws, $f_{ij}^n, j = 1, ..., J$ and knows the vector of sourcing potential, ξ_j, but learns about the firm-specific shock ϵ_j^n to the sourcing potential of country j only after paying the fixed cost to source from

country j. We assume that firms have rational expectations when making their sourcing strategy decision.[22] The expected profits of selecting a sourcing strategy $\mathcal{J}$ for a firm with core efficiency φ and a vector of fixed costs of sourcing f_{ij}^n are given by:

$$\Pi\left(\mathcal{J}, \varphi, f_{ij}^n\right) = \varphi^{\sigma-1} B E_\epsilon \left(\left(\gamma \Theta_i \left(\mathcal{J}, \epsilon \right) \right)^{(\sigma-1)/\theta} \right) - \sum_{j \in \mathcal{J}} f_{ij}^n, \tag{31}$$

In a setting with a large number of countries, the firm faces a very large discrete choice problem to solve for it's optimal sourcing strategy; if there are 65 countries, the firm selects between 2^{65}, which is roughly 10^{19}, possible sourcing strategies. Clearly, calculating the profits of each of these strategies is infeasible. Instead, we will apply an algorithm first developed by Jia (2008) to tractably solve the firm's problem.

This algorithm works as follows. Given a core efficiency φ and a sourcing strategy $\mathcal{J}$, the expected marginal benefit of adding country j to a firm's sourcing strategy is:

$$\varphi^{\sigma-1} \gamma^{(\sigma-1)/\theta} B E_\epsilon \left(\left(\Theta_i \left(\mathcal{J}, \epsilon \right) \right)^{(\sigma-1)/\theta} - \left(\Theta_i \left(\mathcal{J} \setminus j, \epsilon \right) \right)^{(\sigma-1)/\theta} \right) - f_{ij}^n.$$

We define a mapping, $V_j(\mathcal{J})$ which takes a value of one if the marginal benefit of including country j in the sourcing strategy $\mathcal{J}$ is positive, and takes a value of zero otherwise. Because of increasing differences in the profit function, this mapping is an increasing function itself. Jia (2008) showed that when starting from the set $\mathcal{J}^0$ (which contains no country in the sourcing strategy), an iterative application of the V-operator leads to a lower bound of the sourcing strategy, that is the optimal sourcing strategy contains at least the countries contained in this set. Similarly, when starting from the set $\mathcal{J}^1$ (which contains all countries in the sourcing strategy), the iterative application of the V-operator leads to an upper bound for the optimal sourcing strategy. Should the two bounds not overlap, one then only needs to evaluate the profits resulting from all possible combinations between the two bounds. In the presence of a high degree of complementarity, there is the potential for this algorithm to lead to a large number of possible choices between the two bounds, hence rendering this approach infeasible. Intuitively, the iterative process might stall too quickly if it is optimal for firms to add or drop countries from the set $\mathcal{J}$ only in pairs (or larger groups).

Fortunately, in our application, we have found this approach to lead to completely overlapping lower and upper bounds in the vast majority of our simulations and to only a small number of countries differing in the bounds in those cases in which the bounds did not completely overlap. In principle, the algorithm may still be useful even if a sizable number of location sets need to be evaluated; for example, one could assume that the firm evaluates the lower and upper bounds and a random vector of alternative sourcing strategies that are contained in the two bounds.[23]

[22] We calculate the expectation by assuming that the shocks $\log \epsilon_j^n$ are distributed according to a normal distribution with standard deviation equal to the root mean squared error of the OLS regression based on equation (29), and use 100 Halton draws to evaluate the expectation.

[23] In two important aspects the sourcing strategy problem here is simpler than the original problem analyzed by Jia (2008), who studied the location choices of Walmart and Kmart. First, by assuming monopolistic competition we abstract away from strategic interactions between firms. Second, while the firm's choice in her problem manifests

Before describing the estimation method we complete the parameterization of the model. Specifically, we assume that the firm-country-specific fixed cost, f_{ij}^{n}, are drawn from a log-normal distribution with dispersion parameter, β_{disp}^{f} and scale parameter $\log \beta_c^f + \beta_d^f \log \text{distance}_{ij} + \log \beta_l^f \, \text{language}_{ij}$. Hence, fixed cost of sourcing are specified to be a function of distance and common language. We set the fixed cost of domestic sourcing to 0 since all firms in our sample feature positive purchases of domestic inputs. Some firms may be better at screening for foreign suppliers than other firms and thus we assume that the fixed cost draws are rank-correlated across countries, with correlation 0.9. Without rank-correlation, as the number of countries becomes large, it becomes very likely that each firm has a very low fixed cost of sourcing from at least one country, which would imply that the number of importers would go to one.[24] In addition to the fixed cost parameters, we estimate at this step the firm core productivity levels' dispersion parameter, κ, and the scale parameter B, which also includes the CES price index and the mass of firms.

Overall, we are left with the following six parameters to be estimated: $\delta = \left[B, \kappa, \beta_c^f, \beta_d^f, \beta_l^f, \beta_{\text{disp}}^f\right]$.[25] We simulate $S = 50,000$ U.S. firms, that is we draw for each firm a core-productivity shock from a uniform distribution (which, given a parameter guess κ, can be inverted to yield the Pareto distributed firm core productivity level), and a J-dimensional vector of fixed cost shocks from a standard normal distribution (which, given a parameter guess β^f, can be used to calculate the firm-country specific fixed cost level).[26] Note there is no relationship between the number of simulated firms and the number of actual firms in the data. The model assumes that we have a continuum of firms whose core efficiency, fixed cost draws, and country-specific efficiency shocks follow particular distributions, and we use the simulated firms as evaluation points of these distributions.

We use the simulated firms to construct the following three sets of moments, which are directly affected by the level of the fixed costs, β^f, the scale term, B, and the dispersion of core productivity levels, κ:

1. The share of importing firms (about 24 percent in the data). This is simply a scalar, and we label this moment in the data as m_1 and the simulated moment as $\hat{m}_1(\delta)$.

2. The share of firms that sources from a particular country. We label this $J \times 1$ vector of moments in the data as m_2 and the simulated moment vector as $\hat{m}_2(\delta)$.

3. The share of firms that purchases inputs from a country less or more than the q-th percentile of input purchases in the data, where $q = (25, 50, 90)$. We label this $4 \cdot J \times 1$ vector of moments in the data as m_3 and the simulated moment vector as $\hat{m}_3(\delta)$.

We describe the difference between the moments in the data and in the simulated model by

geographic dependence, in contrast, in our setting only the source countries' distances to the U.S. matter, but not the distances between the countries themselves.

[24]We use a Gaussian copula and the marginal distributions of fixed cost for a particular country are distributed lognormal.

[25]Note that we also set $\underline{\varphi}_{\text{US}} = 1$, as it scales input purchases equivalently to an increase in B.

[26]We use Halton draws for the fixed cost shocks, which have better coverage properties than usual pseudo-random draws. See Chapter 6 in Train (2009) for a discussion.

$\hat{y}(\delta)$:

$$\hat{y}(\delta) = m - \hat{m}(\delta) = \begin{bmatrix} m_1 - \hat{m}_1(\delta) \\ m_2 - \hat{m}_2(\delta) \\ m_3 - \hat{m}_3(\delta) \end{bmatrix}$$

The following moment condition is assumed to hold at the true parameter value δ_0:

$$E\left[\hat{y}(\delta_0)\right] = 0 \tag{32}$$

The method of simulated moments selects the model parameters that minimize the following objective function:

$$\hat{\delta} = \arg\min_{\delta} \left[\hat{y}(\delta)\right]^\top \mathbf{W} \left[\hat{y}(\delta)\right], \tag{33}$$

where $\mathbf{W}$ is a weighting matrix. At this point we use as weights simply the identity matrix, but we intend to use the optimal weighting matrix in future iterations of this paper.

The parameter estimates are displayed in Table 9 below.

Table 9: Simulated Method of Moments Estimates

B	0.0001
κ	1.5773
β_c^f	0.0592
β_d^f	0.0027
β_l^f	0.7134
β_{disp}^f	2.2022
Objective value	1.09

Notes: Standard errors yet to be calculated.

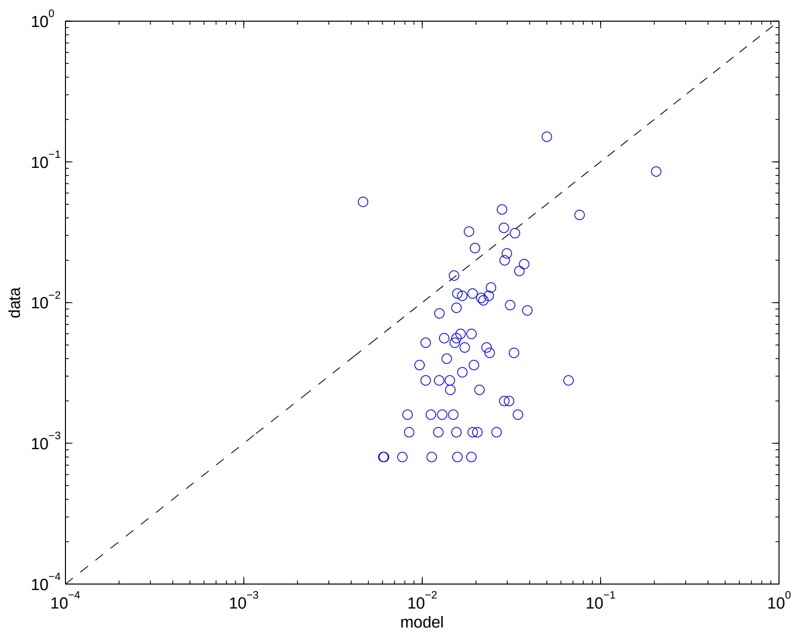
We find that the fixed cost of sourcing are only slightly increasing in distance and that sourcing from countries with a common language reduces the fixed cost by about 30 percent. The fixed cost of sourcing are reasonable in magnitude. We find a median fixed cost estimate ranging from 40,000 to 60,000 USD. However due to the lognormal distribution the fixed cost can be quite large for some individual firm-country combinations. The shape parameter of the Pareto distribution, κ, is lower than the elasticity of substitution, $\sigma - 1$, and hence we need to truncate the support of the core efficiencies at a very large positive number in order to guarantee a finite price index. Other studies also have estimated Pareto size distributions that require truncation (see the discussion in Antràs and Yeaple (2013)). We proceed by describing the fit of the data by the estimated model.

4.4 Fit of the model

Fit of the share of importers In the data around 24 % of US firms import. In the simulated model we slightly under-predict this fraction and predict 21 % importers.

Fit of the share of importers by country Our fit of the share of importers by country is displayed in Figure 3. Overall, the model does a pretty good job at predicting which countries will have a large or small number of importers. We under-predict the number of firms that import from Canada (around 15 percent in the data and around 5 percent in the model) and over-predict the number of firms that import from China (around 20 percent in the model and 9 percent in the data). Hence, in our calibrated model China is the most popular destination country. Overall, it appears that we generally tend to overpredict the share of importers in countries where these shares are small in the data.

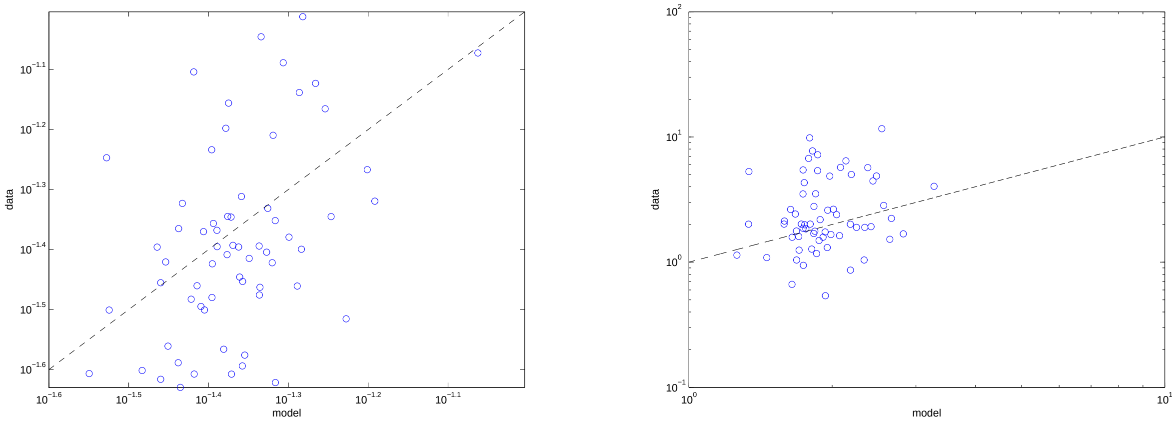
Figure 3: Share of importers by country in model and data



Fit of percentiles by country Our model also provides a reasonably good fit of the percentiles of imports by country. In Figure 4 we display the Median and 90th percentile import purchases by country. The median input purchase from Canada is around 30,000 USD and around 80,000 USD from China in both model and data (all these figures are conditional on buying a positive amount from the country). The 90 percentile purchase by firm from China is around 4 million USD and less than 2 million from Canada - again, the model fits those numbers well. However, the model has a more difficult time fitting the distribution of U.S. input purchases. The median domestic input purchase is 560,000 USD in the data and only 1,000 USD in the estimated model. The 90th percentile of the domestic input purchases is 7.8 million USD in the data and 22,000 USD in the

model.[27] We are currently exploring how to better match the size distribution of domestic input purchases.

Figure 4: Median and 90th percentile imports in model and data



Notes: In logarithmic scales. Line: x=x.

5 Counterfactuals

To be written.

6 Conclusion

To be written.

[27]In the simulated model, a domestic input purchase of 560,000 is reached by the 98th percentile, and a purchase of 7.8 million is reached and exceeded by the top .6 percent of firms.

A Theory Appendix

Proof of Proposition 1

Consider two firms with productivities φ_H and φ_L, with $\varphi_H > \varphi_L$. Denote by $\mathcal{J}_i(\varphi_H) = \{j : I_{ij}(\varphi_H) = 1\}$ and $\mathcal{J}_i(\varphi_L) = \{j : I_{ij}(\varphi_L) = 1\}$ the optimal sourcing strategies of these firms, and suppose that $\mathcal{J}_i(\varphi_H) \neq \mathcal{J}_i(\varphi_L)$ (when $\mathcal{J}_i(\varphi_H) = \mathcal{J}_i(\varphi_L)$ the result in the Proposition holds trivially). For firm φ_H to prefer $\mathcal{J}_i(\varphi_H)$ over $\mathcal{J}_i(\varphi_L)$, we need

$$\varphi_H^{\sigma-1}\left(\gamma\Theta_i\left(\mathcal{J}_i\left(\varphi_H\right)\right)\right)^{(\sigma-1)/\theta} B_i - \sum_{j\in\mathcal{J}_i(\varphi_H)} f_{ij} > \varphi_H^{\sigma-1}\left(\gamma\Theta_i\left(\mathcal{J}_i\left(\varphi_L\right)\right)\right)^{(\sigma-1)/\theta} B_i - \sum_{j\in\mathcal{J}_i(\varphi_L)} f_{ij},$$

while φ_L preferring $\mathcal{J}_i(\varphi_L)$ over $\mathcal{J}_i(\varphi_H)$ requires

$$\varphi_L^{\sigma-1}\left(\gamma\Theta_i\left(\mathcal{J}_i\left(\varphi_H\right)\right)\right)^{(\sigma-1)/\theta} B_i - \sum_{j\in\mathcal{J}_i(\varphi_H)} f_{ij} < \varphi_L^{\sigma-1}\left(\gamma\Theta_i\left(\mathcal{J}_i\left(\varphi_L\right)\right)\right)^{(\sigma-1)/\theta} B_i - \sum_{j\in\mathcal{J}_i(\varphi_L)} f_{ij}.$$

Combining these two conditions, we find

$$\left[\varphi_H^{\sigma-1} - \varphi_L^{\sigma-1}\right]\left[\Theta_i\left(\mathcal{J}_i\left(\varphi_H\right)\right)^{(\sigma-1)/\theta} - \Theta_i\left(\mathcal{J}_i\left(\varphi_L\right)\right)^{(\sigma-1)/\theta}\right]\gamma^{(\sigma-1)/\theta} B_i > 0.$$

For firm φ_H to prefer $\mathcal{J}_i(\varphi_H)$ over $\mathcal{J}_i(\varphi_L)$, we need

$$\varphi_H^{\sigma-1}\left(\gamma\Theta_i\left(\mathcal{J}_i\left(\varphi_H\right)\right)\right)^{(\sigma-1)/\theta} B_i - \sum_{j\in\mathcal{J}_i(\varphi_H)} f_{ij} > \varphi_H^{\sigma-1}\left(\gamma\Theta_i\left(\mathcal{J}_i\left(\varphi_L\right)\right)\right)^{(\sigma-1)/\theta} B_i - \sum_{j\in\mathcal{J}_i(\varphi_L)} f_{ij},$$

while φ_L preferring $\mathcal{J}_i(\varphi_L)$ over $\mathcal{J}_i(\varphi_H)$ requires

$$\varphi_L^{\sigma-1}\left(\gamma\Theta_i\left(\mathcal{J}_i\left(\varphi_H\right)\right)\right)^{(\sigma-1)/\theta} B_i - \sum_{j\in\mathcal{J}_i(\varphi_H)} f_{ij} < \varphi_L^{\sigma-1}\left(\gamma\Theta_i\left(\mathcal{J}_i\left(\varphi_L\right)\right)\right)^{(\sigma-1)/\theta} B_i - \sum_{j\in\mathcal{J}_i(\varphi_L)} f_{ij}.$$

Combining these two conditions, we find

$$\left[\varphi_H^{\sigma-1} - \varphi_L^{\sigma-1}\right]\left[\Theta_i\left(\mathcal{J}_i\left(\varphi_H\right)\right)^{(\sigma-1)/\theta} - \Theta_i\left(\mathcal{J}_i\left(\varphi_L\right)\right)^{(\sigma-1)/\theta}\right]\gamma^{(\sigma-1)/\theta} B_i > 0.$$

Given $\varphi_H > \varphi_L$, this necessarily implies $\Theta_i(\varphi_H) > \Theta_i(\varphi_L)$.

Proof of Proposition 2

As noted in the main text, when $(\sigma-1)/\theta > 1$, the profit function in (10) features increasing differences in (I_{ij}, I_{ik}) for $j, k \in \{1, ..., J\}$ with $j \neq k$. Furthermore, it also features increasing differences in (I_{ij}, φ) for any $j \in J$. Invoking Topkis's monotonicity theorem, we can then conclude that for $\varphi_H \geq \varphi_L$, we must have $(I_{i1}(\varphi_H), I_{i2}(\varphi_H), ..., I_{iJ}(\varphi_H)) \geq (I_{i1}(\varphi_L), I_{i2}(\varphi_L), ..., I_{iJ}(\varphi_L))$. Naturally, this rules out a situation in which $I_{ij}(\varphi_H) = 0$ but $I_{ij}(\varphi_L) = 1$, and thus we can conclude that $\mathcal{J}_i(\varphi_L) \subseteq \mathcal{J}_i(\varphi_H)$ for $\varphi_H \geq \varphi_L$.

Proof of Proposition 3

Given a vector of wages, equations (16) and (17) determine the equilibrium values of B_i and N_i. Notice that the firm-level global sourcing problem depends only on B_i, w_i and exogenous parameters, and not directly

on N_i. As a result, if a unique solution for B_i exists, all thresholds $\tilde{\varphi}_{ij}$ for any pair of countries (i,j) will be pinned down uniquely, given wages. Hence, if a unique solution for B_i in equation (16) exists, we can ensure that there will be a unique value of N_i solving (17). Let us then focus on studying whether (16) indeed delivers a unique solution for B_i.

For given wages, the equilibrium condition (16) can be rearranged as follows

$$w_i f_e = B_i \int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty} \left(\gamma \Theta_i(\varphi)\right)^{(\sigma-1)/\theta} \varphi^{\sigma-1} dG_i(\varphi) - w_i \int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty} \sum_{j\in\mathcal{J}_i(\varphi)} f_{ij} dG_i(\varphi), \tag{34}$$

where remember that $\vartheta(i)$ is defined as $\vartheta(i) = \{j \in J : \tilde{\varphi}_{ij} \leq \tilde{\varphi}_{ik} \text{ for all } k \in J\}$ and thus satisfies

$$\left(\tilde{\varphi}_{i\vartheta(i)}\right)^{\sigma-1} B_i \left(\gamma T_{\vartheta(i)} \left(\tau_{i\vartheta(i)} w_{\vartheta(i)}\right)^{-\theta}\right)^{(\sigma-1)/\theta} = w_i f_{i\vartheta(i)}. \tag{35}$$

Remember also that $\Theta_i(\varphi) \equiv \sum_{k\in\mathcal{J}_i(\varphi)} T_k (\tau_{ik} w_k)^{-\theta}$, and $\mathcal{J}_i(\varphi) \subseteq J$ is the set of countries for which a firm based in i with productivity φ has paid the associated fixed cost of offshoring $w_i f_{ij}$.

Computing the derivative of the right-hand-side of (34) with respect to B_i, and using (35) to eliminate the effects working through changes in $\tilde{\varphi}_{i\vartheta(i)}$, we can write this derivative as simply

$$\int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty} \frac{\partial \left(\varphi^{\sigma-1} \left(\gamma\Theta_i(\varphi)\right)^{(\sigma-1)/\theta} B_i - w_i \sum\limits_{j\in\mathcal{J}_i(\varphi)} f_{ij}\right)}{\partial B_i} dG_i(\varphi) > 0. \tag{36}$$

The fact that this derivative is positive follows directly from the firm's global sourcing problem in 10. In particular, holding constant the firm's sourcing strategy $\mathcal{J}_i(\varphi)$ – and thus $\Theta_i(\varphi)$ –, it is clear that an increase in B_i will increase firm level profits $\varphi^{\sigma-1}(\gamma\Theta_i(\varphi))^{(\sigma-1)/\theta} B_i - w_i \sum_{j\in\mathcal{J}_i(\varphi)} f_{ij}$. Now such an increase in B_i might well affect the profit-maximizing choice of $\mathcal{J}_i(\varphi)$ – and thus $\Theta_i(\varphi)$ –, but firm profits could not possibly be reduced by those changes, since the firm can always decide not to change the global sourcing strategy in light of the higher B_i and still obtain higher profits.[28] We can thus conclude that the right-hand-side of (34) is monotonically increasing in B_i.

It is also clear that when $B_i \to \infty$, all firms will find it optimal to source everywhere and the right-hand-side of (34) becomes

$$B_i \left(\gamma \sum\nolimits_{k\in J} T_k (\tau_{ik} w_k)^{-\theta}\right)^{(\sigma-1)/\theta} \int_{\underline{\varphi}_i}^{\infty} \varphi^{\sigma-1} dG_i(\varphi) - w_i \sum_{j\in J} f_{ij}$$

and thus goes to ∞. Conversely, when $B_i \to 0$, no firm can profitably source to any location, given the positive fixed costs of sourcing, and thus the right-hand-side of (34) goes to 0.

It thus only remains to show that the right-hand-side of (34) is a *continuously* non-decreasing function of B_i. This may not seem immediate because firm-level profits jump discontinuously with B_i when such changes in B_i lead to changes in the global sourcing strategy of firms. It can be shown, however, that

[28] Following the same steps as in the proof of Proposition 1 we can show that both $\Theta_i(\varphi)$ and $\sum_{j\in\mathcal{J}_i(\varphi)} f_{ij}$ are actually non-decreasing in B_i. This result is immaterial for the proof of existence and uniqueness in the case of free entry, but can be used to proof the same result for the case of an exogenous number of firms N_i.

$$\int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty} \frac{\partial\left(\left(\Theta_i\left(\varphi\right)\right)^{(\sigma-1)/\theta} B_i\varphi^{\sigma-1}\right)}{\partial B_i} dG_i\left(\varphi\right)$$

is continuously differentiable in B_i. To see this, one can first follow the same steps as in the proof of Proposition 1 to show that $\Theta_i\left(\varphi; B_i\right)$ must be non-decreasing not only in φ, but also in B_i and $B_i\varphi^{\sigma-1}$. We can then represent $\left(\Theta_i\left(\varphi\right)\right)^{(\sigma-1)/\theta} B_i\varphi^{\sigma-1}$ as a non-decreasing step function in φ, in which the jumps occur at different levels of $B_i\varphi^{\sigma-1}$. This is analogous to writing

$$\left(\Theta_i\left(\varphi\right)\right)^{(\sigma-1)/\theta} B_i\varphi^{\sigma-1} = \begin{cases} \theta_1 B_i\varphi^{\sigma-1} & \text{if } \varphi < a_1/B_i^{1/(\sigma-1)} \\ \theta_2 B_i\varphi^{\sigma-1} & \text{if } a_1/B_i^{1/(\sigma-1)} \leq \varphi < a_2/B_i^{1/(\sigma-1)} \\ \vdots & \vdots \\ \theta_J B_i\varphi^{\sigma-1} & \text{if } a_{J-1}/B_i^{1/(\sigma-1)} \leq \varphi \end{cases}. \tag{37}$$

Hence, we have

$$\begin{aligned}\int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty} \left(\Theta_i\left(\varphi\right)\right)^{(\sigma-1)/\theta} B_i\varphi^{\sigma-1} dG_i\left(\varphi\right) &= \int_{\tilde{\varphi}_{i\vartheta(i)}}^{a_1/B_i^{1/(\sigma-1)}} \theta_1 B_i\varphi^{\sigma-1} dG_i\left(\varphi\right) + \\ &\int_{a_1/B_i^{1/(\sigma-1)}}^{a_2/B_i^{1/(\sigma-1)}} \theta_2 B_i\varphi^{\sigma-1} dG_i\left(\varphi\right) + ... + \int_{a_{J-1}/B_i^{1/(\sigma-1)}}^{\infty} \theta_J B_i\varphi^{\sigma-1} dG_i\left(\varphi\right).\end{aligned}$$

It is then clear that the derivative of this expression with respect to B_i is a sum of continuous functions of B_i, and thus is continuous in B_i itself. Using similar arguments we can next show that

$$\int_{\tilde{\varphi}_{i\vartheta(i)}}^{\infty} \frac{\partial\left(w_i \sum_{j\in\mathcal{J}_i(\varphi)} f_{ij}\right)}{\partial B_i} dG_i\left(\varphi\right) \tag{38}$$

is also continuously differentiable in B_i. First, a simple proof by contradiction can be used to show that $\sum_{j\in\mathcal{J}_i(\varphi)} f_{ij}$ is non-decreasing in $B_i\varphi^{\sigma-1}$. More specifically, suppose that for $\left(B_i\varphi^{\sigma-1}\right)_H > \left(B_i\varphi^{\sigma-1}\right)_L$ we also had $\sum_{j\in\mathcal{J}_{iH}} f_{ij} < \sum_{j\in\mathcal{J}_{iL}} f_{ij}$. Given the non-decreasing dependence of $\Theta_i\left(\varphi\right)$ on $B_i\varphi_i^{\sigma-1}$, we would then have

$$\left(\gamma\Theta_{iH}\left(\varphi\right)\right)^{(\sigma-1)/\theta}\left(B_i\varphi^{\sigma-1}\right)_L - \sum\nolimits_{j\in\mathcal{J}_{iH}} f_{ij} > \left(\gamma\Theta_{iL}\left(\varphi\right)\right)^{(\sigma-1)/\theta}\left(B_i\varphi^{\sigma-1}\right)_L - \sum\nolimits_{j\in\mathcal{J}_{iL}(\varphi)} f_{ij},$$

which clearly contradicts $\mathcal{J}_{iL}$ being optimal given $B_i\varphi^{\sigma-1} = \left(B_i\varphi^{\sigma-1}\right)_L$. With this result, $\sum_{j\in\mathcal{J}_i(\varphi)} f_{ij}$ can then be expressed as a step function analogous to that in (37), in which the position of the steps is continuously differentiable in B_i. This in turn ensures that (34) is continuous in B_i and concludes the proof that there exists a unique B_i that solves equation (16).

Algebra of the Special Case $\sigma - 1 = \theta$

[TO BE INCORPORATED]

B Data Appendix

B.1 Sample

Table B.1 provides details of all firms in the Economic Censuses with positive sales and employment. The first row corresponds to firms that consist only of manufacturing establishments ("M" firms). The second row presents information for all firms with one or more manufacturing establishments and at least one establishment outside of manufacturing ("M+" firms). Together, these two types of firms comprise our sample.

Table B.1: Sample of firms

Firm Type	Firms	Imports $millions	Empl $millions	Sales $billions	Fraction Importers
Manufacturing Only (M)	238,800	75,938	5,866	1,230	0.23
Manufacturing Plus (M+)	11,500	829,594	20,573	9,180	0.77
Other (O)	4,001,000	99,937	77,204	12,553	0.03
Wholesale Only (W)	300,000	240,654	3,484	2,300	0.31
Wholesale and Other (WO)	7,500	141,740	6,426	2,252	0.51
Total	4,561,700	1,387,863	113,553	27,515	0.06

Notes: Table provides information on firms in the Economic Census with positive sales and employment. Analysis in paper based on all M and M+ firms. Numbers rounded for disclosure avoidance. Imports exclude products classified under mining.

References

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

Antràs, Pol, *Global Production: A Contracting Perspective* 2014. in-progress book manuscript available at http://scholar.harvard.edu/antras/scholar.harvard.edu/antras/pages/books.

_ and Elhanan Helpman, "Global Sourcing," *Journal of Political Economy*, 2004, *112* (3).

_ and Stephen Yeaple, "Multinational Firms and International Trade Structure," in "Handbook of International Economics," Vol. 4, Elsevier, 2013.

_ , Luis Garicano, and Esteban Rossi-Hansberg, "Offshoring in a Knowledge Economy," *The Quarterly Journal of Economics*, 2006, *121* (1), 31–77.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, February 2012, *102* (1), 94–130.

_ , Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple, "Innovation and Production in the Global Economy," Working Paper 18972, National Bureau of Economic Research April 2013.

_ , Svetlana Demidova, Peter J Klenow, and Andrés Rodríguez-Clare, "Endogenous Variety and the Gains from Trade," *American Economic Review*, April 2008, *98* (2), 444–450.

Barro, Robert and Jong Wha Lee, "A New Data Set of Educational Attainment in the World, 1950-2010," 2010. NBER Working Paper No. 15902.

Bernard, Andrew B. and Teresa C. Fort, "Factorlyess Goods Producers in the US," Working Paper 19396, NBER June 2013.

_ , Andreas Moxnes, and Karen Helene Ulltveit-Moe, "Two-Sided Heterogeneity and Trade," SSRN Scholarly Paper ID 2341044, Social Science Research Network, Rochester, NY October 2013.

Bernard, Andrew B, J. Bradford Jensen, Stephen J Redding, and Peter K Schott, "Firms in International Trade," *Journal of Economic Perspectives*, August 2007, *21* (3), 105–130.

_ , _ , _ , and _ , "The Margins of US Trade," *American Economic Review*, April 2009, *99* (2), 487–493.

_ , _ , _ , and _ , "Wholesalers and Retailers in US Trade," *American Economic Review*, May 2010, *100* (2), 408–413.

Blaum, Joaquin, Claire Lelarge, and Michael Peters, "Non-Homothetic Import Demand: Firm Productivity and Quality Bias," *Unpublished paper*, 2013.

Blum, Bernardo S, Sebastian Claro, and Ignatius Horstmann, "Facts and Figures on Intermediated Trade," *American Economic Review*, May 2010, *100* (2), 419–423.

Carballo, Jerónimo, Gianmarco I. P. Ottaviano, and Christian Volpe Martincus, "The Buyer Margins of Firms' Exports," CEP Discussion Paper dp1234, Centre for Economic Performance, LSE 2013.

Chaney, Thomas, "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, August 2008, *98* (4), 1707–1721.

Costinot, Arnaud, "An Elementary Theory of Comparative Advantage," *Econometrica*, 2009, *77* (4), 1165–1192.

Dragusanu, Raluca, "Firm-to-Firm Matching Along the Global Supply Chain," 2014. mimeo Harvard University.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

_ , Marcela Eslava, David Jinkins, C. Krizan, and James Tybout, "A Search and Learning Model of Export Dynamics," 2012. mimeo Penn State University.

_ , Samuel Kortum, and Francis Kramarz, "An Anatomy of International Trade: Evidence From French Firms," *Econometrica*, 2011, *79* (5), 1453–1498.

_ , _ , _ , and Raul Sampognaro, "Firm-to-Firm Trade: Imports, Exports, and the Labor Market," 2013. unpublished manuscript.

Ethier, Wilfred J., "National and International Returns to Scale in the Modern Theory of International Trade," *The American Economic Review*, June 1982, *72* (3), 389–405.

Feenstra, Robert C. and J. Bradford Jensen, "Evaluating Estimates of Materials Offshoring from U.S. Manufacturing," *Economic Letters*, 2012, *117*, 170–173.

Fort, Teresa C., "Breaking Up Is Hard To Do: Why Firms Fragment Production Across Locations," Working Paper 13-35, Center for Economic Studies, U.S. Census Bureau 2013.

Garetto, Stefania, "Input Sourcing and Multinational Production," *American Economic Journal: Macroeconomics*, April 2013, *5* (2), 118–151.

Goldberg, Pinelopi Koujianou, Amit Kumar Khandelwal, Nina Pavcnik, and Petia Topalova, "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India," *The Quarterly Journal of Economics*, November 2010, *125* (4), 1727–1767.

Gopinath, Gita and Brent Neiman, "Trade Adjustment and Productivity in Large Crises," *American Economic Review*, 2013, p. Forthcoming.

Grossman, Gene M. and Esteban Rossi-Hansberg, "Trading Tasks: A Simple Theory of Offshoring," *American Economic Review*, 2008, *98* (5), 1978–97.

_ and _ , "Task Trade Between Similar Countries," *Econometrica*, 2012, *80* (2), 593–629.

Grossman, Gene M and Giovanni Maggi, "Diversity and Trade," *American Economic Review*, December 2000, *90* (5), 1255–1275.

Hall, Robert E. and Charles I. Jones, "Why do some countries produce so much more output per worker than others?," *Quarterly Journal of Economics*, 1999, *114*, 83–116.

Halpern, László, Miklós Koren, and Adam Szeidl, "Imported inputs and productivity," *Working Paper, CEU*, 2011, *8*, 28.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein, "Estimating Trade Flows: Trading Partners and Trading Volumes," *The Quarterly Journal of Economics*, May 2008, *123* (2), 441–487.

Jia, Panle, "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry," *Econometrica*, November 2008, *76* (6), 1263–1316.

Johnson, Robert C. and Guillermo Noguera, *Fragmentation and trade in value added over four decades*, Working Paper, Dartmouth College, January 2012.

Krugman, Paul, "Scale Economies, Product Differentiation, and the Pattern of Trade," *The American Economic Review*, December 1980, *70* (5), 950–959.

Loecker, Jan De, Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik, "Prices, Markups and Trade Reform," Working Paper 17925, National Bureau of Economic Research March 2012.

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

Milgrom, Paul and John Roberts, "The Economics of Modern Manufacturing: Technology, Strategy, and Organization," *The American Economic Review*, June 1990, *80* (3), 511–528.

Morales, Eduardo, Gloria Sheu, and Andrés Zahler, "Gravity and Extended Gravity: Using Moment Inequalities to Estimate a Model of Export Entry," Working Paper 19916, National Bureau of Economic Research February 2014.

Olley, G Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–97.

Oostendorp, Remco H., "The Standardized ILO October Inquiry 1983-2003," mimeo, Amsterdam Institute for International Development 2005.

Ramondo, Natalia and Andrés Rodríguez-Clare, "Trade, Multinational Production, and the Gains from Openness," *Journal of Political Economy*, April 2013, *121* (2), 273–322.

Rodríguez-Clare, Andrés, "Offshoring in a Ricardian World," *American Economic Journal: Macroeconomics*, April 2010, *2* (2), 227–258.

Rossi-Hansberg, Esteban and Mark L. J Wright, "Establishment Size Dynamics in the Aggregate Economy," *American Economic Review*, December 2007, *97* (5), 1639–1666.

Tintelnot, Felix, "Global Production with Export Platforms," 2012. mimeo Penn State University.

Train, Kenneth, *Discrete Choice Methods with Simulation*, 2nd ed., Cambridge University Press, 2009.

Vives, Xavier, "Nash equilibrium with strategic complementarities," *Journal of Mathematical Economics*, 1990, *19* (3), 305–321.