Online Appendix: The Modern Wholesaler: Global Sourcing, Domestic Distribution, and Scale Economies

Sharat Ganapati
Georgetown University, CESifo, and NBER
sganapati@gmail.com

A Data Sources and Construction

A.1 Data Used

I bring together a variety of censuses and surveys conducted by the United States Census Bureau, Department of Transportation, and Department of Homeland Security covering international trade, domestic shipments and both the manufacturing and wholesale sectors. I use the Census of Wholesale Trade, Census of Manufacturers, Longitudinal Firm Trade Transaction Database, Commodity Flow Survey, and the Longitudinal Business Database, from 1992 to 2012.

The Census of Wholesale Trade (CWH) collects data every five years on the entire universe of wholesale establishments, subdividing wholesalers by both type and ownership structure. In particular the CWH divides wholesale establishments into merchant wholesalers (MW) and manufacturers sales and branch offices (MSBO). As this paper considers wholesalers that are independent from manufacturers, I exclude MSBO and other similar establishments from analysis. However, aggregate census statistics may not distinguish between these two establishment forms and overestimate the wholesaler market presence. Notably, distribution centers owned by downstream buyers, such as those by large retail chains are systematically excluded from this census.[22] This dataset is central to the analysis and provides administrative data on operating costs, merchandise purchases, total sales, goods sold, and buyer types.[23] Wholesale industries distributing products with sales consisting of more than 50% non-manufactured goods are excluded. This includes certain petrochemical segments distributing crude oil, and all agricultural and mining sectors. Data from 1992 and 2012 are not directly comparable to data from 1997-2002 due to changes in industry classification systems. (The 1992 data uses the Standard Industrial Classifications and 2012 data uses a significant revision of the NAICS system.)

The Census of Manufactures (CMF) aggregates data every five years on the universe of manufacturing establishments. This extensively used dataset provides information on a range of values, including total shipments and various operating and capital expenses. I focus on the value of shipments in producer values. This database helps in calculating the total domestic absorption of manufacturing products as well as the share of goods shipped directly by manufacturers. As with the

[22] The second largest building in the United States by usable space is the Target Import Warehouse in Lacey, Washington. However I assume that such buildings are classified as retailers and not wholesalers, with Target operating as the final destination.

[23] The biggest drawback of this data is the lack of quantity data. I will explicitly account for this in the model and estimates by considering units in terms of producer prices.

CWH, the CMF lacks explicit quantity data for the vast majority of industries (notable exceptions include cement, concrete, and steel).

The Commodity Flow Survey (CFS) is conducted every five years and collects data on a random selection of shipments for a set of establishments. This data is collected for both wholesale and manufacturing establishments and is used to construct crosswalks between manufacturing and wholesale sectoral designations. Additionally the micro-data includes statistics on the origin, destination, and value of individual shipments, as well as export status.

The Longitudinal Firm Trade Transaction Database (LFTTD) tracks and links imports and exports by product at the firm level. This database catalogues all import and export transactions by date from 1992 onwards in terms of both value and quantity. Tying all the datasets together, the Longitudinal Business Database provides a way to link individual establishments from the CWH, CMF, and CFS at the firm level, as well as linking these firms with trade data from the LFTTD. The process of merging these databases and further details are reported below.

A.2 Census of Wholesale Trade (CWH)

The U.S. Census Defines a wholesaler in the 2007 North American Industry Classification System (NAICS) as:

> The Wholesale Trade sector comprises establishments engaged in wholesaling merchandise, generally without transformation, and rendering services incidental to the sale of merchandise. The merchandise described in this sector includes the outputs of agriculture, mining, manufacturing, and certain information industries, such as publishing.
>
> The wholesaling process is an intermediate step in the distribution of merchandise. Wholesalers are organized to sell or arrange the purchase or sale of (a) goods for resale (i.e., goods sold to other wholesalers or retailers), (b) capital or durable non-consumer goods, and (c) raw and intermediate materials and supplies used in production.
>
> Wholesalers sell merchandise to other businesses and normally operate from a warehouse or office. These warehouses and offices are characterized by having little or no display of merchandise. In addition, neither the design nor the location of the premises is intended to solicit walk-in traffic. Wholesalers do not normally use advertising directed to the general public. Customers are generally reached initially via telephone, in-person marketing, or by specialized advertising that may include Internet and other electronic means. Follow-up orders are either vendor-initiated or client-initiated, generally based on previous sales, and typically exhibit strong ties between sellers and buyers. In fact, transactions are often conducted between wholesalers and clients that have long-standing business relationships.
>
> This sector comprises two main types of wholesalers: merchant wholesalers that sell goods on their own account and business to business electronic markets, agents, and brokers that arrange sales and purchases for others generally for a commission or fee.

I focus on the first type of business, merchant wholesalers, which are further described as:

> Merchant wholesale establishments typically maintain their own warehouse, where they receive and handle goods for their customers. Goods are generally sold without transformation, but may include integral functions, such as sorting, packaging, labeling, and other marketing services.

In addition, I omit three types of wholesalers, first those that are classified as Manufacturer's Sales and Branch Offices (MSBO), those that are classified as own-brand importers and markets, and firms classified as agents/electronic markets. This specifically excludes what Bernard and Fort (2015); Bernard et al. (2017) consider former manufacturers that may have transitioned from domestic manufacturing into foreign manufacturing and domestic distribution. If these firms are included as wholesalers, the wholesale shares of distribution increase more dramatically.

For clarity, I've reproduced the selected portions of the Economic Census form from 2007 for NAICS 423190 - Electrical Goods Wholesalers in Figure A1 (forms from 1997 and 2002 are similar are publicly available). In question 19, I exclude firms that are classified as "14: Own-brand importer and marketer", "20: Manufacturers' sales branch or office", "41-48: Agent, broker, or commission merchant", "49: Electronic market", or "77: Other broker or agent".

Wholesalers are classified according to their NAICS code. A market is defined as all downstream buyers that buy and sell from these NAICS codes. For example, Code 421610 refers to wholesalers participating in the resale of "Electrical Apparatus and Equipment, Wiring Supplies and Construction Material". While establishments may appear to belong to multiple codes, this project only considers the Census-designated code. Future research projects may further explore multiple-industry wholesalers. Firms may own establishments in multiple NAICS wholesale sectors. I divide foreign imports proportionally between sectors, weighting by the volume of goods purchased.

Sales are aggregated considering the wholesaler's purchase cost from their upstream source, net of export sales, and correcting for inventory adjustments. Prices are in manufacturers' dollars and computed using the ratio between the sales to downstream buyers divided by upstream purchases by the wholesalers. Wholesale industries that derive more than 50% of revenues from products that are not manufactured are removed from analysis. These industries pertain primarily to mining and agricultural products. Additionally, NAICS sectors 424710 and 424720 dealing with petroleum and petroleum products are removed, as are NAICS sectors 424810, 424820, and 424940 that deal with beer, wine, and tobacco products. Petroleum products are removed as a result of the industry taking a unique form due to the ownership and distribution of pipeline networks. Alcohol and tobacco products are often regulated at the wholesaler level by individual states. Some states do not allow for direct sourcing by downstream retailers and force the usage of wholesalers, rendering my model of wholesaling spurious.

A.2.1 Wholesaler Prices

Wholesaler prices are systematically denoted in producer prices. Therefore a wholesaler price of \$1.3 implies that it costs \$1.3 to indirectly buy \$1 manufactured output (at the "factory gate").

Wholesalers prices p_w are constructed as follows:

$$p_w = \frac{\tilde{p}_w q_w}{\tilde{p}_m q_m},$$

where $\tilde{p}_m$ and $\tilde{p}_w$ represent the price paid by the wholesaler to a manufacturer and the price paid by a downstream firm to a wholesaler respectively. Variable q_m represents the quantity purchased from a manufacturer, and q_w represents the quantity sold by a wholesaler. In practice, quantity data is unavailable for most industries, so $p_m q_m$ is approximated by

$$C_m = p_m q_m,$$

where C_m represents the expenditures of a wholesaler on manufactured goods. Similarly

$$R_w = \tilde{p}_w q_w,$$

where R_w represents the revenue of a wholesaler. In Figure A1, C_m corresponds to question 16(b) and R_m corresponds to question 5(a).

I clean the data so wholesaler inventory changes are netted out, thus:

$$p_w = \frac{\tilde{p}_w}{\tilde{p}_m}.$$

As estimation requires a normalization, I set $\tilde{p}_m = 1$, so wholesaler prices p_w are all relative to producer prices $\tilde{p}_m$. I explore robustness to this price definition in Appendix B.3, where I allow differentiated buyers to face different wholesaler prices.

In addition, I require operating cost data to derive accounting markups, this corresponds to question 16(b) in Figure A1.

A.2.2 Wholesaler Sales Data

Wholesaler sales data is broken down by product origin by merging the LFTTD and CWH on firm-level characteristics. First, total sales are derived from the line item referring to "Sales and operating receipts." Purchases from manufacturers are derived from the line referring to "Purchases of merchandise for resale."

Data from the LFTTD denotes the imports by country of origin. Countries (outside of the U.S.) are divided into two varieties using the World Bank's World Development Indicators Database from 1997. Sources with per-capita gross domestic product (GDP) over \$10,000 are categorized as high-income sources. Sources with per-capita GDP under \$10,000 are classified as low-income sources. The cut-off county in my database is Slovenia; all richer countries are high-income sources. Due to extensive literature highlighting the pass-through nature of Hong Kong's economy (Feenstra and Hanson (2004)), imports from Hong Kong and Macau are re-classified as Chinese imports.

As the World Bank estimates are not complete, I manually categorize a small subset of countries. Afghanistan, Iraq, Kosovo, Myanmar, Nauru, Sao Tome and Principe, South Sudan, Somalia, and Timor-Leste are classified as low income countries. San Marino is classified as a high income country.

Figure A1: Selected Survey Questions: 2007 Economic Census

U.S. DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. CENSUS BUREAU
FORM
WH-42319 (12/07/2006)

2007 ECONOMIC CENSUS
Electrical Goods

OMB No. 0607-0929: Approval Expires 12/31/2008

DUE DATE
FEBRUARY 12, 2008

Mail your completed form to:
U.S. CENSUS BUREAU
1201 East 10th Street
Jeffersonville, IN 47134-0001

Please read the accompanying information sheet(s) before answering the questions.

Need help or have questions about filling out this form?

WH-42319

. . .

5 SALES, SHIPMENTS, RECEIPTS, OR REVENUE

		Mark "X" if None	2007 $ Bil.	Mil.	Thou.	Dol.
A. Sales and operating receipts *(Include the gross selling value of business conducted for others. Include shipping and handling charges. Exclude sales taxes and Hawaii's General Excise Tax.)*	0100	☐				

. . .

16 SELECTED EXPENSES

		Mark "X" if None	2007 $ Bil.	Mil.	Thou.	Dol.
A. Operating expenses *(Include payroll. Exclude cost of goods sold and interest expense.)*	0140	☐				
B. Purchases of merchandise for resale, net of returns, allowances, and trade and cash discounts *(Include amounts allowed for trade-ins.)*	1160	☐				

. . .

19 **B.** TYPE OF OPERATION
Which ONE of the following best describes this establishment's principal type of operation in 2007?
(Mark "X" only ONE box.)

Merchant wholesaler, buying and selling on own account

0600

- 12 ☐ Importer
- 13 ☐ Exporter
- 11 ☐ Merchant wholesale distributor or jobber
- 14 ☐ Own-brand importer and marketer
- 20 ☐ **Manufacturers' sales branch or office**

Agent, broker, or commission merchant

0600

- 41 ☐ Auction company
- 42 ☐ Broker, representing buyers and sellers
- 43 ☐ Commission merchant
- 44 ☐ Import agent
- 45 ☐ Export agent
- 46 ☐ Manufacturers' agent
- 49 ☐ **Electronic market - business-to-business marketplace that facilitates the sale of goods via the Internet or other electronic means, and operates on a commission or fee basis**

42319053

Overseas territories of the UK, Netherlands, and France are classified according to their parent country's status (see Gibraltar, Curacao, and St. Martin/Sint Maarten).

Wholesaler purchases of domestic manufactured goods are computed by subtracting imports from total merchandise purchases for resale. Finally, sales are adjusted to only consider domestic buyers. I subtract the percentage of sales and purchases that are used for export shipments. This export data is collected directly on the CWH forms. Additionally there are a subset of wholesaler firms that participate in multiple NAICS wholesale sectors. I allocate imports proportionally by the cost of goods sold between the multiple sectors.

A.3 Outside Share (Direct Sourcing) Data Construction

Both the summary statistics in Section I and the estimation routine in Section III, require the construction of the total downstream market size and the share of the downstream market not served by U.S. based wholesalers (the outside option). As wholesalers in the Census of Wholesale Firms (CWH) and and manufacturing producers in Census of Manufacturers (CMF) use different classification systems, a series of NAICS Wholesale to NAICS Manufacturers code concordances are used. See Ganapati (2015) for an overview of the process. In addition, the Import-Export Database (LFTTD) uses the Harmonized System (HS) of good classification, and the Commodity Flow Survey (CFS) uses the Standardized Classification of Transported Goods (SCTG). Ganapati (2015) also uses the micro-data in the CFS and the LFTTD to provide concordances between the various NAICS, HS and SCTG codes at different levels of aggregation.

Total domestic absorption is computed as:

$$\begin{aligned}\text{Total Domestic Absorption} &= \text{Domestic Production} \\ &+ \text{International Imports} \\ &- \text{International Exports.}\end{aligned}$$

Data on domestic production originates from the CMF as the sum of all domestically manufactured products. Data on international imports and exports originates from the LFTTD. For domestic wholesalers in the LFTTD, values are deflated by average wholesaler markups over manufacturer prices. This produces "total domestic absorption" in terms of producer's prices. Since manufacturers and producers are not modeled in this paper, these prices are considered fixed. Alternative computation uses the CFS for domestic production and international export data. All prices for manufactured goods are deflated by the BEA series for "Chain-Type Price Indexes for Materials Inputs"

Domestic absorption by wholesalers is computed as:

$$\begin{aligned}\text{Domestic Wholsaler Absorption} &= \text{Domestically Sourced Wholesaler Shipments} \\ &+ \text{Wholesaler Imports} \\ &- \text{Wholesaler International Exports.}\end{aligned}$$

Table A1: **Number of Source Countries and Products by Market Share Quantile**

Multi-location Firms by Quantile			
Share	Year		
Quantile	1997	2002	2007
0-10	0.1	0.1	0.1
10-20	0.1	0.1	0.2
20-30	0.1	0.2	0.2
30-40	0.2	0.3	0.3
40-50	0.3	0.3	0.4
50-60	0.3	0.4	0.5
60-70	0.5	0.5	0.7
70-80	0.6	0.8	0.8
80-90	0.9	1.1	1.3
90-99	2.0	2.4	2.7
99-99.5	5.1	6.3	6.5
99.5+	9.9	12.4	13.6

Multi-location Firms by Quantile			
Share	Year		
Quantile	1997	2002	2007
0-10	0.2	0.3	0.4
10-20	0.4	0.4	0.5
20-30	0.5	0.8	0.9
30-40	0.7	1.1	1.4
40-50	1.0	1.4	1.7
50-60	1.4	1.8	2.6
60-70	1.9	2.5	3.5
70-80	5.0	4.1	5.0
80-90	5.0	8.8	11.6
90-99	13.7	18.0	24.6
99-99.5	54.1	77.0	73.4
99.5+	137.4	183.6	213.8

The first two components are computed using the combination of the CWH along with the LFTTD. The CWH reports total shipments and total exports, the LFTTD reports the total imports of a firm. Wholesaler international exports are computed using the self-reported CWH figure for total exports, alternatively the LFTTD may also be used.

A.4 Detailed Wholesaler Statistics

Tables A1-A3 highlight additional wholesaler statistics by wholesaler size rank. These are aggregates across all wholesalers.

A.5 Distribution of Buyer Types

Data on the mass of buyer types m_j comes from the Commodity Flow Survey, combing purchases from wholesalers and manufacturers. Product codes (SCTG classifications) from wholesaler shippers (with NAICS Codes) are used to convert shipments from manufacturer NAICS codes to wholesaler NAICS codes.

I present an additional fact that describes the time evolution of buyer types in the Commodity Flow Survey.

Fact 7 *The distribution of buyer types has slightly skewed towards larger shipments over time.*

One hypothesis explaining the shift towards wholesaling is the spread of "just in time" manufacturing and supply practices. These business models forgo a small number of large deliveries for a larger number of smaller shipments. This provides downstream buyers with more flexibility and reduces inventory costs. In aggregate, such practices would imply that there is a shift towards smaller order sizes. If wholesalers are more adept at shipping smaller orders, then this may induce

Table A2: **Market Shares and Import Probabilities by Market Share Quantile**

Multi-location Firms by Quantile			
Share Quantile	Year 1997	2002	2007
0-10	0.0001%	0.0001%	0.0001%
10-20	0.0003%	0.0003%	0.0003%
20-30	0.0006%	0.0006%	0.0005%
30-40	0.0010%	0.0010%	0.0009%
40-50	0.0015%	0.0015%	0.0013%
50-60	0.0023%	0.0023%	0.0021%
60-70	0.0036%	0.0035%	0.0033%
70-80	0.0059%	0.0059%	0.0057%
80-90	0.0114%	0.0115%	0.0114%
90-99	0.0404%	0.0426%	0.0461%
99-99.5	0.1740%	0.1970%	0.2356%
99.5+	0.8241%	1.0197%	1.1335%

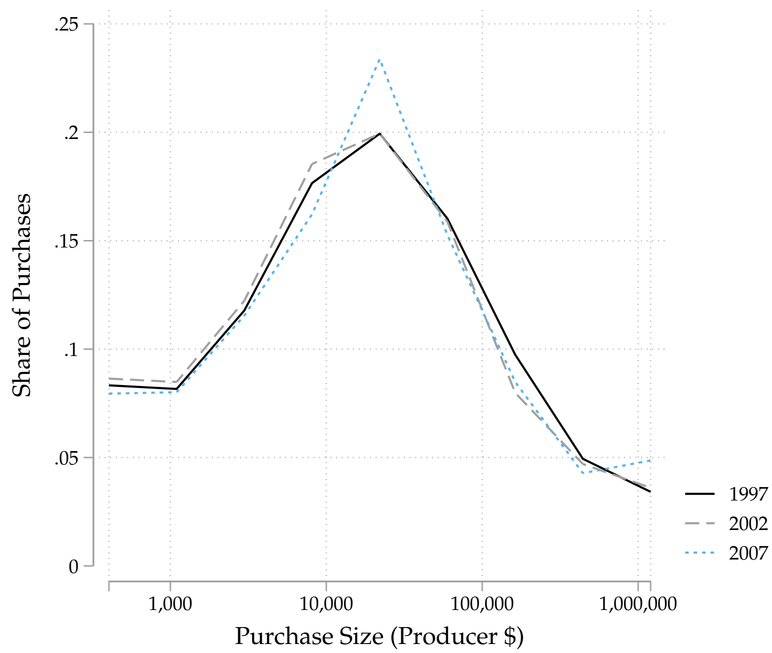
Multi-location Firms by Quantile			
Share Quantile	Year 1997	2002	2007
0-10	5%	6%	8%
10-20	6%	9%	10%
20-30	9%	10%	13%
30-40	11%	13%	16%
40-50	13%	16%	19%
50-60	15%	18%	22%
60-70	19%	22%	26%
70-80	23%	26%	30%
80-90	27%	31%	36%
90-99	39%	42%	48%
99-99.5	60%	62%	67%
99.5+	74%	78%	81%

Table A3: **Number of Locations by Market Share Quantile**

Multi-location Firms by Quantile			
Share Quantile	Year 1997	2002	2007
0-10	0%	0%	0%
10-20	0%	0%	0%
20-30	0%	0%	1%
30-40	1%	1%	1%
40-50	1%	1%	2%
50-60	2%	2%	3%
60-70	4%	4%	4%
70-80	7%	7%	7%
80-90	13%	13%	14%
90-99	28%	30%	31%
99-99.5	50%	53%	57%
99.5+	63%	68%	71%

Multi-location Firms by Quantile			
Share Quantile	Year 1997	2002	2007
0-10	1.0	1.0	1.0
10-20	1.0	1.0	1.0
20-30	1.0	1.0	1.0
30-40	1.0	1.0	1.0
40-50	1.0	1.0	1.0
50-60	1.0	1.0	1.0
60-70	1.0	1.1	1.1
70-80	1.1	1.1	1.1
80-90	1.2	1.2	1.3
90-99	1.8	2.0	2.1
99-99.5	4.7	5.9	6.9
99.5+	14.2	20.7	23.9

Figure A2: **Distribution of Buyers**



Figures in real 2007 dollars.

a shift of buyers switching to wholesalers. However, this has not occurred, as shown in Figure A2; Downstream buyers have slightly increased the average size of their orders over time.[24]

A.6 Geographic Differentiation

In lieu of a continuous distance measure, this project discretizes downstream buyer location by U.S. state[25], which are each located in 4 regions and 9 divisions. This project considers three distinct levels of distance with regards to the downstream buyer: wholesalers that are located in the same state, wholesalers located in the same census division and wholesalers located in a different census division. Figure A3 displays these divisions.

An alternative approach that would allow for tractable computation would be to map distance directly to distance indicator variables. This would prevent issues arising from considering the distance between New York and Connecticut differently than the distance between New York and New Jersey, due to Census division classifications. Instead of considering buyers that are within the same census division or region, the alternative would be to consider other states within pre-specified distance bands. For example, distance band 1 for New York would include all wholesalers in states that are reachable within 4 hours (250 miles) and distance band 2 would include all wholesalers in states that are within 8 hours (500 miles). Preliminary results show that estimates in Sections

[24]A related fact shows that the geographic distribution of buyers has not significantly changed over the same time period.

[25]The District of Columbia is redefined as a state for this project.

Figure A3: **U.S. Census Regions and Divisions**

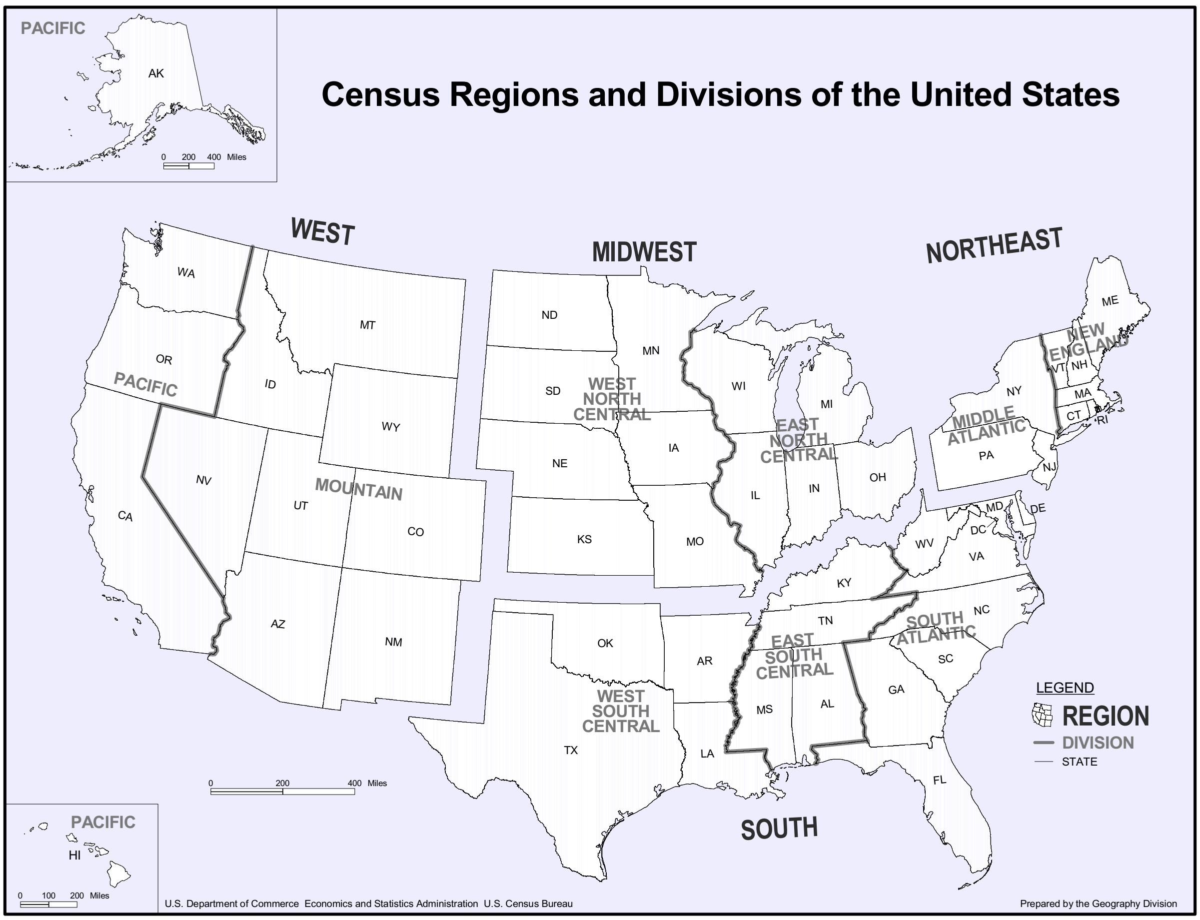
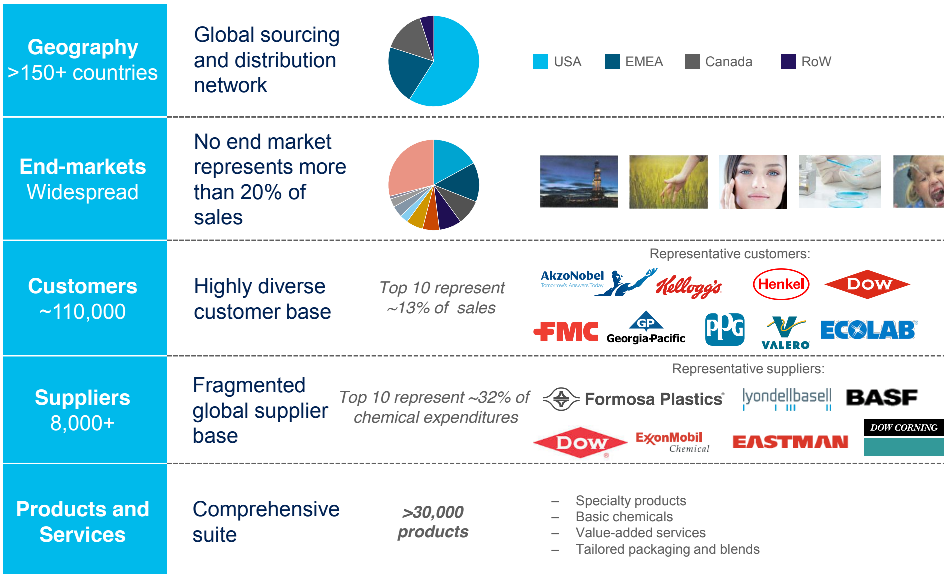


Figure A4: **Univar Presentation at 2015 Barclays Industrial Distribution Forum**



Source: Univar Investor Relations

III and IV are largely consistent and the aggregate estimates in Section V are similar. However, the geographic breakdown is slightly changed, with the surplus gains due to intermediation slightly rising in small New England and South Atlantic States (in particular Rhode Island and Delaware) and slightly falling in rural Mountain States (Wyoming and Montana). The primary restriction here is the lack of computing power, enabling full estimation.

A.7 Wholesaler Case Study

Consider the case of specialty industrial chemicals. This sector grew 28% between 2008 and 2013; however, the share of products distributed by independent wholesalers increased 37%. Industry reports (Elser et al., 2010; Jung et al., 2013, 2014) highlight two types of observations, (a) why particular downstream buyers contract with wholesalers instead of manufacturers and (b) what differentiates successful wholesalers from unsuccessful wholesalers.

Downstream buyers face heterogenous barriers to directly purchasing chemicals from a manufacturer. According to a 2009 Boston Consulting Group survey, 80% of downstream buyers with purchases valued under €100,000 sourced goods indirectly through wholesalers, while larger purchasers nearly always sourced directly from a manufacturer. Downstream buyers value traditional distributor attributes such as price, quality, and globally sourced varieties, and are differentiated on two characteristics, their size and geographic location.[26]

In the industrial chemical market, wholesaler distributors perform three functions as they (a) source products from multiple manufacturers, (b) repackage these products, and (c) ship these

[26] Smaller downstream buyers "typically lack the critical mass needed to tap into low-cost sources for chemicals from China, Eastern Europe, or the Middle East." In addition, these downstream buyers not only value price, product quality, and technical support, they prize flexibility and speed of delivery, which are highly correlated with geographic proximity.

products to downstream buyers. While the global market for distributors is still fragmented, it is experiencing rapid consolidation, with the three largest companies in 2011 holding 39% of the North American market. In particular, the largest distributors have grown faster than the market, driven by both organic expansion and market acquisitions. In contrast, smaller distributors face increasing fixed costs, as they try to "combine global reach with strong local presence." (Jung et al., 2013)

Consider one of the large speciality chemical distributors, Univar. A slide detailing their business plan is presented in Figure A4. Univar is a large industrial chemical wholesaler with North American shipments of approximately $10.4 billion in 2014. The company was formed in 1928, increasing its distribution footprint through acquisitions and expansions. Today, it sources 30,000 varieties of chemicals and plastics from over 8,000 internationally distributed suppliers. Univar uses its 8,000 employees to run a distribution network spanning hundreds of locations to supply 111,000 buyers.[27]

Downstream buyers may need a variety of chemicals, and they may source these chemicals directly from manufacturers such as DuPoint and BASF, or indirectly through Univar. However, BASF and DuPont facilities may be located in distant locations and only stock their own product lines. Instead of individually sourcing chemicals, downstream buyers may pay a markup and have Univar do this for them, where Univar would source the shipments from each respective chemical manufacturer and reship them to a convenient loading bay. This tradeoff between convenience and price is one of the central dynamics underpinning the wholesale industry. This also offers insight into why the wholesale industry may be gaining market share, as the proliferation of new global sources and varieties may make it harder to optimally source intermediate products for production.[28]

B Demand Systems

This section provides micro-foundations for the indirect downstream profit functions used in Section II. This provides support for both the two-stage demand system and allows for simple extensions. While this specific toy demand model provides micro-foundations for the exact demand structure presented in the main paper's model, it is slightly generalizable, while still providing the needed structure. There are two critical elements, the first requiring a single-input invertible production function, and the second requiring that the expectation of the marginal cost is sufficient for the wholesaler's decision in the last demand stage (in period t_4).[29]

B.1 Downstream Profit Maximization (1st Demand Stage)

To highlight downstream buyers' choices of purchase quantity before the realization of idiosyncratic match shocks, consider a hypothetical downstream buyer. Assume that these downstream buyers produce output using a single input, such that output $q = x$, where q is the single input. Downstream

[27]Univar's business plan is summarized in a slide presented as Appendix Figure A4.

[28]Feenstra and Weinstein (2017) show that the number of manufactured varieties in the U.S. has increased over time due to global trade.

[29]The logic here closely follows Hausman et al. (1995), switching the buyer's problem to consider a producer's profit maximization.

buyers face constant elasticity of substitution (CES) demand for $x > 0$ units, with elasticity $\sigma > 1$ and demand-shifter $\eta > 0$. Additionally, suppose there are fixed cost of production f drawn from some distribution $F(\cdot)$.

First, I solve the firm's problem disregarding the fixed cost. Demand takes the form:

$$x = \eta p^{-\sigma}$$

Under such a CES demand framework, these downstream buyers charge markup μ, which is a function of the elasticity of substitution σ:

$$\mu = \frac{\sigma}{\sigma - 1}.$$

This markup is invariant of the demand shifter η. The optimal price, p^*, charged by such a downstream buyer is the product of the marginal cost of production mc and the markup μ:

$$p^* = mc \cdot \mu.$$

This price can be plugged back into the demand equation, solving for the optimal q^*:

$$x^* = \eta\left(\mu \cdot mc\right)^{-\sigma}.$$

Since the production function is one-to-one with the input, $q^* = x^*$. However, this assumes that downstream buyer marginal cost mc is known. In the two-stage decision, downstream buyers must choose q^{**} in a first period, with knowledge of only the possible distribution of mc. Then in the second period, downstream buyers choose p^{**} to clear the market. Solving through backwards induction, conditional on x^{**}, a downstream buyer chooses p^{**} such that:

$$p^{**} = \left(\frac{x^{**}}{\eta}\right)^{-1/\sigma}$$

Then in the first stage, a wholesaler solves:

$$\max \mathbb{E}\left[\left(p\left(x\right) - mc\right) \times x\right]$$

Plugging in values, iterating expectations of marginal cost, and taking first order conditions:

$$\begin{aligned}
\pi\left(x\right) &= x\left(\frac{x}{\eta}\right)^{-1/\sigma} - x\mathbb{E}\left[mc\right] \\
\pi'\left(x\right) &= \frac{\sigma - 1}{\sigma}\left(\frac{x}{\eta}\right)^{-1/\sigma} - \mathbb{E}\left[mc\right]
\end{aligned}$$

Setting the first order conditions to zero and solving for x^{**}:

$$\begin{aligned}
x^{**} &= \eta\left(\mathbb{E}\left[mc\right]\mu\right)^{-\sigma}. \\
&= q^{**}
\end{aligned}$$

Where the last equality comes from the linear production function. This two-stage demand provides for the same prices and quantities as before while allowing for uncertainty in the realized marginal cost.

If the demand shifter η comes from some underlying distribution $N(\cdot)$, then the distribution of q^* will come from this same distribution scaled by $(\mu \cdot mc)^{-\sigma}$.

Revisiting fixed cost f, expected profits are:

$$E(\pi) = E((p^{**} - mc)\, q^{**}) - f \quad = \quad \tilde{\pi}(E(mc)) - f$$

Where $\tilde{\pi}$ is an increasing function in terms of the expected marginal cost. Production only occurs if $\tilde{\pi} - f > 0$.

Aggregate downstream profits are a decreasing function of marginal cost, thus a reduction in marginal costs increases downstream profits.[30] The second stage's demand decision involves choosing the optimal wholesaler to reduce this marginal cost. Additionally, these profits are a function of the fixed cost f; lowered marginal costs imply that more firms will be able to enter the market. Aggregating across the draws for downstream demand η and the fixed costs f, this produces a mass of buyers M_q that demand q units. If $\mathbb{E}(mc)$ falls, then the mass of M_q will shift upwards. In my model $\mathbb{E}(mc)$ is directly related to $\mathbb{E}(\bar{U})$, the expected utility of indirectly sourcing from a wholesaler.

B.2 Downstream Cost Minimization (2nd Demand Stage)

The indirect downstream profit function can be micro-founded through a simple cost minimization function for a downstream buyer. Suppose the cost of directly sourcing q units is:

$$C_{direct} = qp_0 F(q)$$

Where p_0 is the per-unit cost and $F(q)$ is the per-unit overhead cost of setting up purchases for q units. Suppose the indirect cost of sourcing q units is:

$$C_{indirect} = qp_1$$

Where p_1 is the per-unit cost. For simplicity, suppose there isn't an overhead cost. The logarithm of per-unit costs are then:

$$\begin{aligned} \log\left(\frac{C_{direct}}{q}\right) &= \log(p_0) + \log\left(\frac{F(q)}{q}\right) \\ \log\left(\frac{C_{indirect}}{q}\right) &= \log(p_1) \end{aligned}$$

As long as downstream profits or utility are a function of the difference in per-unit costs, then the estimating equation is appropriate.

[30] Note that $\sigma > 1$.

B.3 Quantity discounts

Business to business transactions often take a form where the sale price is a function of the the quantity purchased. While the estimated model does not directly account for this, a simple modification allows for quantity discounts to be easily added without changing the implication of the model. Suppose that wholesaler price p depends on the purchased quantity q through discount factor $d(q)$ and a mean price p:

$$p_q = p \times d(q).$$

The discount function $d(q)$ is a schedule that multiplies some baseline price conditional on the purchase quantity q.

Simplifying the mean utility δ_q from equation (9) for any wholesaler selling to a buyer purchasing q units produces:

$$\delta_q = \alpha \log p_q + f(q) + \xi$$

Where $f(q)$ represents the different preferences for wholesalers depending on purchase quantity q. Substituting the function for price:

$$U_q = \alpha \log p + \underbrace{\alpha \log d(q) + f(q)}_{\tilde{f}(q)} + \xi$$

Instead of recovering $f(q)$, estimation now recovers $\tilde{f}(q)$. In terms of buyer surplus calculations and market entry estimates, results are essentially unchanged. In terms of marginal cost estimates, similar logic prevails, and this paper computes a mean marginal cost with industry-year fixed effects netting out buyer compositional changes. However for counterfactuals, I assume that this discount structure $d(q)$, through $\tilde{f}(q)$, is invariant. That is prices p_q can only change through p and not through $d(q)$, which will remain fixed.

B.4 Constant Elasticity of Substitution

The choice between wholesalers is modeled as a discrete choice decision and is micro-founded above. This modeling assumption is used both for tractability and realism, even though the majority of international trade research uses a constant elasticity of substitution demand system. However, there is a nice link between CES demand systems and the discrete-choice logit demand systems, as first described by Anderson et al. (1992) and elaborated by De Loecker (2011).

Assume that downstream product demand takes the form:

$$D(p) = \left(\frac{p}{P}\right)^{-\rho} \xi \frac{Y}{P} = (p)^{-\rho} \xi \frac{Y}{P^{1-\rho}}$$

Where Y is total spending, ξ is a demand shifter, ρ is the elasticity of substitution, and the price index P takes the form:

$$P = \left(\int \xi p^{1-\rho} \right)^{\frac{1}{1-\rho}}$$

Wholesaler profit maximization takes the following form:

$$\pi = \max_{p} \left(p - c \right) D\left(p \right),$$

which p denoting the price and c denoting wholesaler marginal cost. Assuming Nash-in-prices competition, the optimization is as follows:

$$\begin{aligned} D\left(p\right) &= -\left(p-c\right) D'(p) = \sigma \frac{(p-c)}{p} D(p) \\ p &= c\frac{\rho}{(\rho - 1)} \end{aligned}$$

So then higher/lower prices due to ξ only operate through its correlation to c. Then wholesaler revenues R are:

$$R = (p)^{1-\rho} \xi \frac{Y}{P^{1-\rho}}$$

Taking a log transform of the wholesaler revenue function produces the relationship:

$$\log R = (1-\rho) \log p + \log \xi + \log \frac{Y}{P^{1-\rho}} \tag{12}$$

Now since revenues are related to market share s and total market size Y as $R = sY$, equation (12) can be rewritten as:

$$\log s = (1-\rho) \log p + \log \xi - \log P^{1-\sigma}$$

This estimating equation is almost identical to the logit estimating equation, with $\alpha = (1-\rho)$. The difference between these models, as noted by Anderson et al. (1992), is clearly in the economic interpretation, but the use of log prices forces identical substitution patterns. Note this model is not directly used in the empirical application, rather I use an aggregation of a nested logit framework. Further work can shows that this is equivalent to a two-level nested-CES demand aggregated across a variety of heterogenous downstream buyers. Both the two-level nested structure of demand and the heterogenous downstream buyers produce substantially more complex aggregate substitution patterns between wholesalers allowing much richer analysis. Critically, the difference between my model and most international trade papers is on the supply-side. Firms do not compete monopolistically, they are allowed to exert variable market power.

B.5 Demand Estimation

B.5.1 Bresnahan et al. (1997) Demand Structure

Following McFadden (1980) and Bresnahan et al. (1997), I assume the distribution of the vector of $\overrightarrow{\epsilon}$ for a given buyer (i, j) is drawn from a "principals of differentiation" (PD) nested logit model.

Formally ϵ is drawn from the distribution $F\left(\overrightarrow{\epsilon}_{i,j}\right) = exp\left(-G\left(e^{-\epsilon_{i,j,0}}, ..., e^{-\epsilon_{i,j,W,I}}\right)\right)$, where $G\left(\cdot\right)$

takes the functional form:

$$G\left(e^{\vec{\delta}_{i,j}}\right) = e^{\delta_{i,j,0}} + \alpha_o \left[\sum_{i\in\mathcal{I}}\left(\sum_{w\in\mathcal{S}_i} e^{\delta_{w.i,j,}/(1-\sigma_i)}\right)^{1-\sigma_i}\right] + \alpha_n \left[\sum_{n\in\mathcal{S}_n}\left(\sum_{w\in\mathcal{W}} e^{\delta_{w,i,j}/(1-\sigma_n)}\right)^{1-\sigma_n}\right]$$

where weights $\alpha_i = \sigma_i / (\sigma_i + \sigma_n)$ and $\alpha_n = 1 - \alpha_i$. Set $\mathcal{S}_i$ includes all wholesaler-source combinations of of variety i and set $\mathcal{S}_n$ includes all wholesaler-source combinations from wholesaler classification n, which correspond to different types of multi-variety wholesalers. The parameter $\sigma = (\sigma_i, \sigma_n)$ must lie inside the unit circle. As either σ goes to zero, the corresponding weight goes to zero, rendering that dimension of product differentiation irrelevant. The first σ_i denotes the correlation of ϵ between direct sourcing, indirectly sourcing from high-income foreign countries, and indirectly sourcing from low-income foreign countries. This allows for products sourced from abroad to be imperfect substitutes for domestically sourced products. The second σ_n denotes the correlation of ϵ of multi-source and single-source wholesalers. This allows for domestic products sourced by globalized wholesalers to be imperfect substitutes for products sourced by domestic-only wholesalers.

B.5.2 Addressing market size

While the structure above allows for significant market segmentation, administrative dataset may still highly limited in available attributes. Markups are heavily reliant on market definitions. In practice equation 4 requires parameter ψ. Empirically this works as a wholesaler-variety shifter $\psi_{w,i}$

$$s^{\psi}_{w,i|j} = \frac{\exp\left(\delta_{w,i|j}\right)}{\exp\left(\delta_{w,i|j}\right) + \psi_{w,i}\sum_{w',i'\neq w,i}\exp\left(\delta_{w,i|j}\right)}.$$

The coefficient $\psi_{w,i}$ is implicitly defined as

$$\exp\left(\delta_{w,i|j}\right) + \psi_{w,i}\sum_{w',i'\neq w,i}\exp\left(\delta_{w,i|j}\right) = \psi\sum_{w,i}\exp\left(\delta_{w,i|j}\right).$$

Identification of this term leverages administrative data. Without this term, estimation of all other demand side parameters is marginally changed.

This step is crucial for matching aggregate data on accounting markups from Table (1). A typical wholesale NAICS code has 4,000 firms. Even with the nesting structure and segmented geographies, market concentration is minimal (see Table (2)), with average HHI measures only increasing from 65.5 to 104.7. With low concentration, competition will realize markups as a function of the demand elasticity and not of competition. To reconcile the accounting markups and concentration data in the underlying data and a model without time-varying demand elasticities, along with the broad nature of NAICS codes, markets must be segmented, using ψ. This parameter simply is the proportion of firms that must compete against each other to rationalize changes in accounting markups over time. As the level of markups without variable market power is pinned down by α, this moment helps pin down effective market size ψ from the changes in markups over time.

Alternatively we could do away with cost shifters to identify the price elasticity and simply use

Table A4: Monte Carlo Simulation for ψ

	Simulation Type			
	1	2	3	4
	Full Data	Single Market S=1	Multiple Markets S	Estimated S
Panel A: $N = 500$, $S = 10$, $\psi = 0.10$				
Markups	2.61	2.96	2.62	2.60
Markup Error (%)		13.57	0.43	-0.11
Markets (S)	10.0	1.0	10.0	10.4
Implied ψ	0.1	1.0	0.1	0.1
Panel B: $N = 500$, $S = 1$, $\psi = 1.0$				
Markups	2.96	2.96	2.96	2.96
Markup Error (%)		0.00	0.00	0.00
Markets (S)	1.0	1.0	1.0	1.0
Implied ψ	1.0	1.0	1.0	1.0

accounting markups to estimate α, as is done in international trade and macroeconomics. But this would produce time-varying estimates of α and complicate measurements of downstream valuation of quality over time.

To illustrate the importance of this step, I conduct a series of Monte Carlo simulations with a simplified set up. I simulate N firms. I see the price, a cost-shifter, and the total sales of all N firms. In addition, I see the aggregate, sales-weighted markup in the market.[31] Crucially, the econometrician may not see how many markets S there are, or which firms belong to which market. I have consumer valuation $\delta_i = -\alpha \cdot p$, where p is drawn from a log-normal distribution with mean 0 and $\sigma = 3$ and where the consumer distaste for price is $\alpha = 3$.

In Table (A4), I conduct four types of simulations and standard IC regressions over 50 runs. First, I assume complete knowledge of which firms are in which markets S. Second I assume that all firms compete in the same market and that $S = 1$. Third I assume I know how many markets S there are. Fourth, I estimate $S = 1/\psi$ using the method above, minimizing the difference between the implied markup and the observed aggregate markup. I do this in Panel A with $N = 500$ and $S = 10$ and then in Panel B with $N = 500$ and $S = 1$. Panel A shows the importance of using ψ when I lack precise data on the makeup of market or market segments. In column two, markups are off by 14%. In column three and four, markups are within 1%. In the last column, I relatively accurately recover ψ and S. Panel B, shows that if this facet is not important, then my estimation will still recover the true parameters.

[31] This is a simplification from the paper, where I assume that I can only see the aggregate markup change over time, however the same logic carries though.

B.5.3 Discrete Choice Estimation Routine

Estimation follows a Generalized Method of Moments technique in the vein of Petrin (2002) and matches both aggregate national market shares and moments derived from the micro-level data.[32]

Assuming away buyer heterogeneity and allowing for one level of nests (the full model follows Bresnahan et al. (1997) and allows for two non-nested levels of nests), I can derive the standard Berry (1994) estimation equation for the relative market share of wholesaler w, selling variety i, that belongs to product nest n:

$$\log s_{w,i} / \log s_0 = \delta_{w,i} + \sigma_n \log s_{w,i|n}, \tag{13}$$

where s_0 represents the share of the outside option, sourcing directly from a manufacturer.[33]

With buyer heterogeneity, the aggregate market share equation is more elaborate:

$$\log s_{w,i} = \log \sum_{j \in \mathcal{J}} \left[s_{0|j} \cdot s^{\sigma}_{w,i|j,n} \cdot \exp\left(\delta_{w,i,j} / \left(1 - \sigma_n\right)\right)\right] b_j \tag{14}$$

Variable $s_{0|j}$ represents the share of direct sourcing from manufacturing by buyers of type j, and $s_{w,i|j,n}$ represents the conditional share of a wholesaler w selling variety i in nest n to customer j. With downstream buyer heterogeneity, alongside wholesaler heterogeneity (that is different wholesalers serve different markets), the demand system provides for flexible substitution patterns and greater variety in markups.

In practice the estimation uses a finite number of buyer types j, each with overall mass b_j. Mean utility $\delta_{w,i,j}$ can be decomposed $\delta_{w,i,j} = \delta_{w,i} + \tilde{\delta}_{w,i,j}$. The first component is common across all downstream buyers and the second is specific to downstream buyers of type j. Solving for $\xi_{w,i}$, equation (14) is operationalized with one level of nests as:

$$\begin{aligned} \xi_{w,i} = \log s_{w,i} - \log \sum_{j \in \mathcal{J}} \left[s_{0,j}\left(\vec{\delta}\right) \cdot s^{\sigma}_{w,i|j,n}\left(\vec{\delta}\right) \cdot \exp\left(\frac{\tilde{\delta}_{w,i,j}}{1-\sigma}\right)\right] b_j \\ - \left(\alpha \log p_{w,i} + \beta_q \log q_j + \sum_{l \in \{state, region\}} \beta_l \mathbb{I}_{l_w = l_d} + \mathbf{x}_{w,i}\alpha_{\mathbf{x}}\right) \end{aligned} \tag{15}$$

This defines a contraction mapping from $\mathbb{R}^N \to \mathbb{R}^N$. By recursively solving for $\xi_{w,i}$, I can solve this system of equations. Multiple levels of nests simply generalize this setup. Unlike the most general form in equation (14), the vector of parameters for unobservable coefficients is set such that $\beta_j = \beta$ for all $j \in J$.

In practice, this contraction mapping is the lengthiest step, as it is difficult to parallelize and requires weeks-long processing time in the confidential census computing cluster. Alternative computation methods such as Mathematical Programming with Equilibrium Constraints (MPEC) are

[32] Estimation proceeds sequentially, starting with demand estimation before moving to estimating the marginal cost and market entry parameters.

[33] If I assume that the unobserved parts of $\delta_{w,n}$ are mean zero, I can run a linear regression and recover $\xi_{w,n}$. However, this means that a wholesaler based in New York will face the same demand in California as in New York, thus the model without buyer heterogeneity is a baseline for the full model.

similarly slow as they require equality constraints for all 600,000 firms to be individually computed and checked.

Aggregates Shares Using observed market shares, a candidate parameter estimate θ, observed prices p, and downstream market characteristics, estimation computes $\xi_{w,i}(\theta)$ for each wholesaler. As shown in Section III, $\xi_{w,i}$ is uncorrelated with a series of instruments z, so the identifying restriction is

$$E\left(\xi_{w,i} z_{w,i}\right) = 0$$

whose empirical analogue is $Z'\xi(\theta)$, where observations are stacked by wholesaler. This set of assumptions will serve to pin down the price coefficient α and substitution σ.

Micro-Level Moments To pin down the coefficients for quantities and geographic indicators, estimation uses a series of moments that use estimated data and compares them with various facets of the survey data. In particular, the estimation routine matches the shares of within metro-area, within state, and within Census region wholesale shipments along with wholesale shipment shares by shipment size.

Large purchases tend to be sourced directly from manufactures and small purchases tend to be sourced indirectly through wholesalers. This is identified using the overall wholesaler market share for a given quantity q:

$$s_{W|q} = \sum_{w\in\mathcal{W}}\sum_{i\in\mathcal{I}}\sum_{j\in J} s_{w,i|j} b_j \mathbb{I}\left\{q_j = q\right\},$$

where $s_{W|q}$ denotes the total market share of all wholesalers conditional on buyer purchase size q. This is a function of observable market share $s_{w,i|j}$ and buyer weights m_j. Additionally, $\mathcal{W}$ represents the set of all wholesalers, $\mathcal{I}$ represents the set of wholesaler varieties, and $\mathcal{J}$ represents the set of buyer types j . Data on b_j in equation (3) comes from the Commodity Flow Survey, which details the share of purchases by location and quantity.

The desirability of a local wholesaler versus a distant wholesaler is identified by the observed share of local, regional, and national shipments:

$$s_{W|l} = \sum_{w\in\mathcal{W}}\sum_{i\in\mathcal{I}}\sum_{j\in J} s_{w,i|j} b_j \mathbb{I}\left\{l_j = l_w\right\}$$

This identifies shipments that do not cross state or regional lines, where the location of the buyer and the location of the wholesaler correspond.

I denote the vector of moments produced by the data as $\mathbf{m}_{data}$ and the estimated moments as $\mathbf{m}(\theta)$.

Correlation Coefficients Estimation uses instruments to identify the nested logit correlation parameters σ. Buyers have similar preferences, but different choice sets, due to regional variations in wholesaler networks. Following the logic of Berry et al. (1995), a wholesaler's entry choices are made before quality $\xi_{w,i}$ is drawn, allowing the number and attributes of competitors to identify σ. In practice, if there are many (few) wholesalers, then within observed wholesaler market shares will be small (large). The intuition is illustrated in a simplified case without observable downstream buyer heterogeneity and one nest. The demand share equation takes the form:

$$\ln\left(s_{w,i}\right)-\ln\left(s_0\right)=\alpha\log p_{w,i}+\sigma\ln\left(s_{w,i|i}\right)+\xi_{w,i}.$$

The market shares of a wholesaler w selling variety i, conditional on selling variety i is denoted $s_{w,i|i}$. This share is correlated with $\xi_{w,i}$ as wholesalers with higher quality draws will not only have higher unconditional market shares, but higher market shares conditional on their attributes. The market of share of direct sourcing from a manufacturer is s_0. A valid instrument needs to satisfy the exogeneity criterion, but at the same time relate to the regressor of interest. As the number and attributes of wholesalers are chosen before the realization of ξ, exogeneity is mechanically satisfied. Estimation generalizes this to include the number of wholesalers with the same sourcing strategy (single-source or multiple-source) and sourcing particular varieties at the regional and state level. I collect these instruments as Z_2.

Moment Function Estimation obtains the parameter estimate $\hat{\theta}$ from minimizing the following criterion equation:

$$\hat{\theta}=\arg_\theta\min G\left(\theta\right)'WG\left(\theta\right), \tag{16}$$

where

$$G\left(\theta\right)=\begin{bmatrix} Z'\xi\left(\theta\right) \\ \mathbf{m}_{data}-\mathbf{m}\left(\theta\right)\end{bmatrix}$$

and W is a weighting matrix. First stage identification uses the identity matrix. But in a two-step procedure, estimation is iterated with the weighting matrix taking the form $W_2=\left[G\left(\hat{\theta}_1\right)G\left(\hat{\theta}_1\right)'\right]^{-1}$ with $\hat{\theta}_1$ denoting the estimates obtained using the identity weighting matrix.

By using the relation, $\delta_{w,i}\left(\sigma\right)=\alpha\log p_{w,i}+x_{w,i}\beta_{\mathbf{x}}+\xi_{w,i}$, estimation can be simplified. Thus conditional on σ, the GMM routine can use the estimation:

$$\hat{\beta}_{IV}\left(\cdot\right)=\left(X'Z\Phi Z'X\right)^{-1}\left(X'Z\Phi Z'X\right)^{-1}\delta_w\left(\sigma,\psi,\beta_l,\beta_q\right)$$

Then I can use a GMM estimator to find σ, ψ, α_l ,and α_q that minimize:

$$J_w\left(\sigma,\psi,\beta_l,\beta_q\right)=\left[\delta_w\left(\sigma,\psi,\beta_l,\beta_q\right)-x\alpha_w\left(\sigma,\psi,\beta_l,\beta_q\right)\right]'Z\phi Z'\left[\delta_w\left(\sigma,\psi,\beta_l,\beta_q\right)-x\beta_{\mathbf{x}}\left(\sigma,\psi,\beta_l,\beta_q\right)\right].$$

B.5.4 Demand Estimation

Formally, I identify the demand parameters α,β,ψ and σ using a modification of Berry and Haile (2014). Define X as the set of attributes defined in the first-stage of the entry game, before the realization of wholesaler quality ξ. This means that a wholesaler has chosen whether they will participate in globalized trade, and what dimension their domestic geographic footprint takes. Define

Table A5: **Downstream Firm Choice (2nd Demand Stage)**

	(1) OLS	(2) Partial IV (Price)	(3) Partial IV (σ)	(4) Full IV
log (Price)	-.153	-2.019	-.423	-1.791
	(0.0038)	(0.0197)	(0.0048)	(0.0203)
σ_i (Varieties)	.92	.851	.76	.694
	(0.0006)	(0.001)	(0.0018)	(0.0018)
Controls	Number of Varieties, Number of Warehouses			
Fixed Effects	Year $\times$ Variety			

Notes: Robust standard errors presented. Columns (1)-(4) show the results without localized market power, nor downstream firm heterogeneity. Columns (1) and (2) omit instruments for $\log(price)$. Column (1) and (3) omit instruments for σ. See text for full regression specification.

Z as a set of variables that shift marginal cost, but not downstream buyer valuations of wholesaler products. Define $M(\alpha, \beta, \psi, \sigma)$ as a set of aggregate moments, such as the predicted share of local wholesale shipments, and where M_d is the observed realization of these moments. I make the following assumptions:

Assumption 1 *For every parameter* $(\alpha, \beta, \psi, \sigma)$ *there is at most one vector* ξ *such that* $s_{w,i}(\xi_{w,i}, \alpha, \beta, \psi, \sigma) - s^0_{w,i} = 0$ *for all* $(w, i) \in \mathcal{W} \times \mathcal{I}$*.*

Assumption 2 $E[\xi_{w,i}|Z, X] = 0$ *for each* $(w, i) \in \mathcal{W} \times \mathcal{I}$

Assumption 3 $E[M(\alpha, \beta, \psi, \sigma) - M_d] = 0$

These assumptions are standard from Berry et al. (1995) and Petrin (2002); a demand invertibility condition, an instrumental variable condition, and a set of aggregate moments. The first condition allows us to invert the observed market shares, conditional on X, and obtain mean valuation $\delta_{w,i}$ for each wholesaler-variety combination $w, i \in \mathcal{W}$.

Assumptions 1, 2, and 3, along with the the structure imposed from the model and set of regularity conditions, identify $\xi_{w,i}$ with probability 1 and the function $s_{w,i}(\cdot)$ is identified. Formally, even without assuming a functional form for $s_{w,i}(\cdot)$, demand identification stems from a modification of Berry and Haile (2014) to allow for aggregate moments.

C Demand Robustness

Table A5 reports results from the estimation of a simplified model of downstream buyer choices from Equation 16. The single nest coefficient σ relates to the substitutability between internationally sourced and domestically sourced goods. Columns (1)-(4) present results from a simplified model without observable buyer heterogeneity and are estimated without the use of the aggregate moments. They are presented with and without appropriate instruments to highlight the importance of controlling for endogeneity. Column (1) omits buyer heterogeneity and neither instruments the wholesaler price nor the correlation coefficient σ. Column (2) instruments for just wholesaler prices and column (3) instruments for just the nest coefficient. Column (4) instruments for both wholesaler prices and the nest coefficient σ.

Columns (1) and (3) do not instrument for wholesaler prices. While downstream buyers appear to value low margins, buyer demand is inelastic. There is a weak relationship between higher prices and lowered sales. This is extremely odd as wholesaling appears to be a low-margin and extremely competitive industry. Instrumenting for wholesaler margins, as in columns (2) and (4), produce much larger (in absolute terms) coefficients and imply that wholesalers all face elastic buyer demand.

The nest coefficient σ relates to the substitutability between internationally sourced and domestically sourced goods. A value of 1 implies zero substitutability between these two categories and a value of 0 implies no differentiation in the substitutability between categories. Without instrumentation, this term will be biased towards 1, as within-type shares will be highly correlated with with total-market shares. This bias is evident in specification (1) and (2), but not in specification (3) and (4).

C.1 Demand Robustness

I consider two further robustness exercises regarding my demand specification; (a) I compress and expand my multi-level nested logit specification and (b) I consider parameter heterogeneity across product-markets. In general, I find that results are largely unchanged.

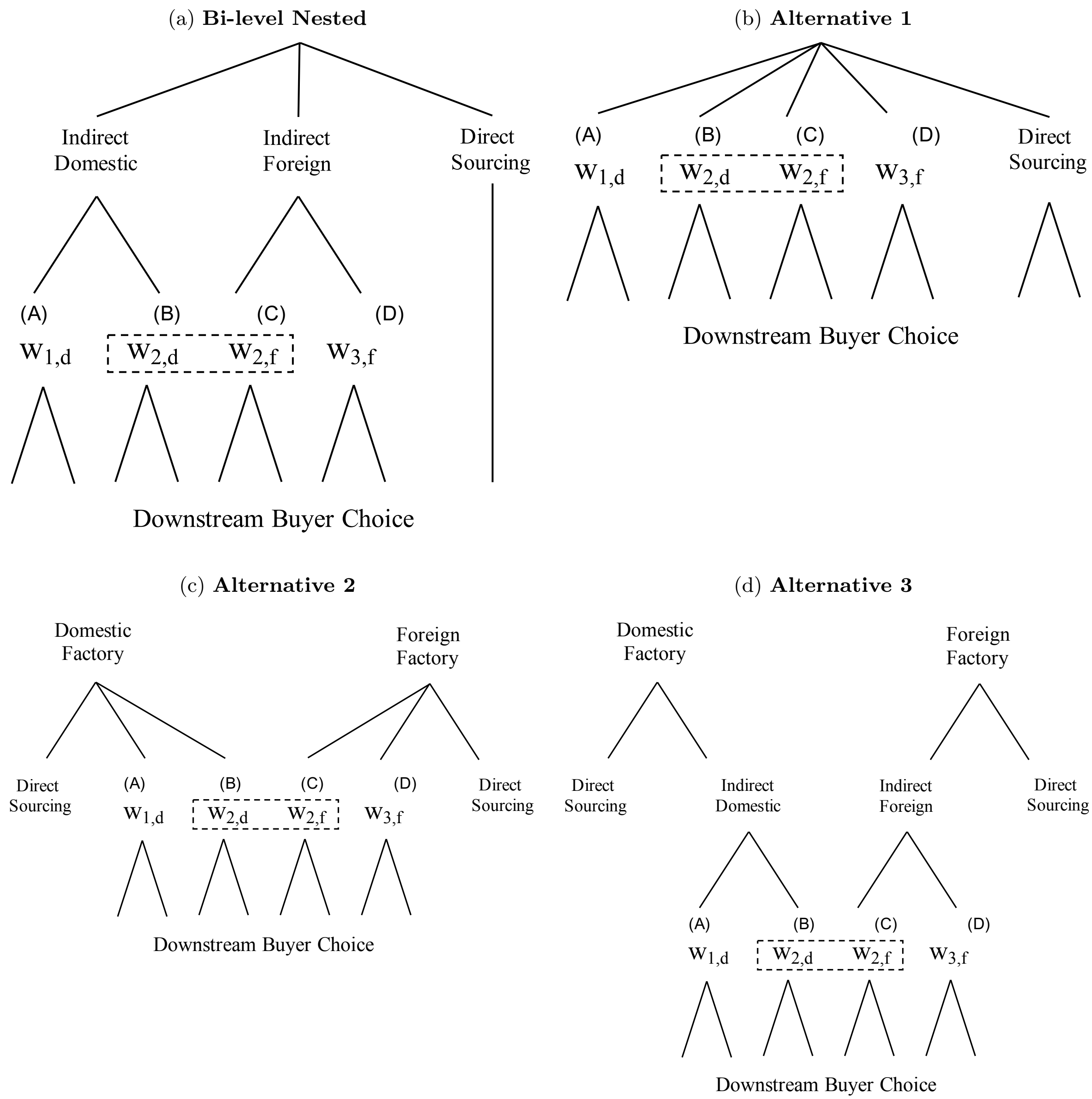
Table A6: **Single-Level Logit Downstream Firm Choice Estimates**

	est/se		est/se		est/se
log (price)	-2.507	Within State Shipment	3.335	log {Shipment Size}	-.314
	(0.023)		(0.145)		(0.054)
log (# Warehouses)	.197	Within Region Shipment	1.356	International Operations	.075
	(0.005)		(0.253)		(0.004)
σ	.636	South Imports × log (HS lines)	.695	North Imports × log (HS lines)	.73
	(0.055)		(0.01)		(0.009)
Fixed Effects	Market × Source, Year × Source				

Notes: Results from optimizing generalized method of moments (GMM) routine using a gradient search. Robust GMM standard errors presented. See text for full regression specification. North refers to high-income country sources. South refers to low-income country sources.

Multi-level Logit Demand In Figure A5, I show a series of alternative nesting patterns for the error term ϵ. Panel (a) shows a classic nested bi-level logit, simplifying the approach in Goldberg (1995). The downside of this model is it implies the substitution between wholesaler types is stronger than between sourcing patterns, which the model in the main paper avoids. Panel (b) compresses the top nesting structure into the second nest. This implies that foreign-sourced products sold by

Figure A5: **Downstream Buyer Sourcing Choice Trees**



Notes: (A) refers to wholesalers that only source from domestic manufacturers. (B) and (C) refer to wholesalers that buy from both domestic and foreign sources, where (B) refers to their domestic purchases and (C) refers to their foreign purchases. (D) refers to wholesalers that only source from abroad. The full model allows for two different types of foreign sources, those from high-income countries and from low-income countries. Additionally, all direct sourcing in lumped together in an outside option.

multi-source wholesalers are similarly substitutable between foreign-sourced products sold by single-source wholesalers and domestically-sourced products sold by multi-source wholesalers. Estimates from such a model are shown in Table A6. In general, this simplified model produces estimates slightly different from the baseline model, as the coefficient estimates α change to rationalize the data to difference in σ. I omit estimation of ψ in this example.

Future projects could further explore the nesting structure in Panels (b) and (c). However, this would require better data on the direct import-share of manufactured goods not at the national level, but at the local (state) level. This variation on the state-level import shares would help identify the substitution parameter σ_{direct} that would govern the top-most nesting structure. This current project aggregates all direct imports at the national level for a data-driven reason. The used import data often lists only the port of landing, not the final destination of an imported product. (As a hypothetical, a disproportionate number of auto parts land in New Jersey, relative to the share auto plants located in the state.) Further work and assumptions are required allocate this import data to downstream users.

Table A7: **Industry-Level Downstream Firm Choice Estimates**

	(1) mean/sd/sem	(2) mean/sd/sem	(3) mean/sd/sem
log (Price)	-1.58	-2.89	-1.45
	[3.66]	[5.93]	[3.75]
	(0.49)	(0.79)	(0.50)
σ_i (Varieties)	0.87		0.89
	[0.40]		[0.44]
	(0.05)		(0.06)
σ_n (Wholesaler Breadth)		0.51	0.81
		[0.34]	[0.70]
		(0.05)	(0.09)
Controls	Number of Varieties, Number of Warehouses		
Fixed Effects	Year × Variety		
Markets	56		

Notes: Results from a 2-stage least squares routine. Robust standard errors presented.

Parameter Heterogeneity In Table A7 I repeat the estimation of my model within each of my 56 product-markets. I use 2-stage least squares estimation, but generalize away from buyer heterogeneity. This produces 56 estimates for the parameter vector (α, σ). I report the average of three critical values for my model and markup calculations, the price coefficient (α^p), and the two parameters governing substitution between nests (σ_i and σ_n).

D Markup Calculations

For simplicity in this Appendix, I assume one level of nests and derive markups when wholesalers exert market power. In terms of notation, $Q_{w,i}$ denotes total sales by wholesaler w selling product i, $s_{w,i|j}$ is the market share conditional on downstream buyer type j, $s_{w,i|j,i}$ is the share conditional on sourcing the same variety i from a different wholesaler, b_j is the mass of downstream buyer type j, and $p_{w,i}$ is the wholesaler's price. Parameters α^p and σ are recovered from demand estimation, and respectively reflect the price sensitivity and substitution elasticities.

I first differentiate the total market size with respect to the wholesaler margin:

$$\begin{aligned}\frac{\partial Q_{w,i}(\mathbf{p})}{\partial p_{w,i}} &= \sum_j \left[\frac{\partial s_{w,i|j}(\mathbf{p})}{\partial p_{w,i}} b_j\right] \\ &= \frac{\alpha^p}{p_{w,i}} \underbrace{\sum_j b_j s_{w,i|j} \left[\frac{1}{1-\sigma}\left[1-\sigma s_{w,i|j,i}-(1-\sigma)s_{w,i|j}\right]\right]}_{\mathfrak{s}_{w,i}} = \frac{\alpha^p}{p_{w,i}}\mathfrak{s}_{w,i}\end{aligned}$$

The new variable $\mathfrak{s}_{w,i}$ summarizes the portion of the demand elasticity that does not directly use any pricing-related terms.

Marginal cost $c_{w,i}$ are as follows for a single product wholesaler:

$$\begin{aligned} c_{w,i} &= p_{w,i} + Q_{w,i}\left(\frac{\partial Q_{w,i}}{\partial p_{w,i}}\right)^{-1} \\ c^*_{w,i} &= p_{w,i} + Q_{w,i}\frac{p_{w,i}}{\alpha^p \mathfrak{s}_{w,i}} = p_w \underbrace{\left(1+\frac{Q_{w,i}}{\alpha \mathfrak{s}_{w,i}}\right)}_{1/\mu_{w,i}}\end{aligned}$$

I denote multiplicative markups as $\mu_{w,i}$.

For a multi-product wholesaler, the price set for varieties i can also have implications for the sales of varieties i' where $i \neq i'$:

$$\begin{aligned}\frac{\partial Q_{w,i'}(\mathbf{p})}{\partial p_{w,i}} &= \sum_j \left[\frac{\partial s_{w,i'|j}(\mathbf{p})}{\partial p_{w,i}} b_j + s_{w,i'|j}(\mathbf{p})\frac{\partial b_j(\mathbf{p})}{\partial p_{w,i}}\right] \\ &= \frac{\alpha^p}{p_{w,i}} \underbrace{(-1)\sum_j b_j s_{w,i'|j} s_{w,i|j}}_{\mathfrak{s}_{i',i}} = \frac{\alpha}{p_{w,i}}\mathfrak{s}_{i',i}\end{aligned}$$

For a multi-product wholesaler selling varieties $i_1, i_2, ...$, consider the matrix of partial derivatives of sales of each sold with respect to to the prices of both the same product and other products sold:

$$\mathbf{\Delta} = \begin{bmatrix} \frac{\partial Q_{w,i_1}}{\partial p_{w,i_1}} & \frac{\partial Q_{w,i_2}}{\partial p_{w,i_1}} & \cdots \\ \frac{\partial Q_{w,i_1}}{\partial p_{w,i_2}} & \frac{\partial Q_{w,i_2}}{\partial p_{w,i_2}} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} = \alpha \begin{bmatrix} \mathfrak{s}_{i_1,i_1} & \mathfrak{s}_{i_2,i_1} & \cdots \\ \mathfrak{s}_{i1,i_2} & \mathfrak{s}_{i_2,i_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} 1/p_{w,i_1} & 0 & \cdots \\ 0 & 1/p_{w,i_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

Solving the system of first order conditions implies that costs are:

$$\begin{pmatrix} c_{w,i_1} \\ c_{w,i_2} \\ \vdots \end{pmatrix} = \begin{pmatrix} p_{w,i_1} \\ p_{w,i_2} \\ \vdots \end{pmatrix} + \Delta^{-1} \begin{pmatrix} Q_{w,i_1} \\ Q_{w,i_2} \\ \vdots \end{pmatrix}$$

D.1 Linking the Model to Data: Multi-Variety Wholesalers

I use a two-step estimator. The underlying data only provides true prices for wholesalers that source a single variety. Prices for multi-variety wholesalers are reported in aggregate. Estimation first recovers cost parameters γ from single-source wholesalers, then recovers marginal costs for all wholesalers. I then re-run estimation across all firms.

The underlying data only provides prices for wholesalers that source a single variety. Prices for multi-variety wholesalers are reported in aggregate. In demand estimation, the instrumental variable strategy can recover price elasticities α, solving the error-in-variables issue for prices. Using summing restrictions, I recover parameters for multi-variety wholesalers that source both domestically and from abroad. This is a product-side interpretation of the logic underpinning De Loecker et al. (2016).

For exposition, assume a wholesaler sells both a domestic variety D and a international variety F. Instead of observing prices $p_{w,F}$ and $p_{w,D}$ separately, I observe the sales weighted average $\bar{p}_w$, where the weights are the known shares, $M_{w,F}$ and $M_{w,D}$. The pricing estimation stage recovers multiplicative markups $\mu_{w,F}$ and $\mu_{w,D}$, as well as data on single-variety wholesalers on $c_w(\cdot)$.

Generalizing away from downstream buyer heterogeneity, this produces the following relations governing prices and costs[34]:

$$\bar{p}_w = M_{w,D} p_{w,D} + M_{w,F} p_{w,F} \tag{17}$$

$$p_{w,D} = \mu_{w,D} c_{w,D} \tag{18}$$

$$p_{w,F} = \mu_{w,F} c_{w,F}. \tag{19}$$

To close the system, I assume that the unobserved component of cost $\nu_{w,i}$ is identical across domestically and internationally sourced goods, rewriting equation (6) as:

$$\log c_{w,F} - \log c_{w,D} = \tilde{\mathbf{x}}_{w,F}\gamma_F - \tilde{\mathbf{x}}_{w,D}\gamma_D \tag{20}$$

This is justified as wholesalers appear to provide the same levels of customer service to their downstream buyers, even if variety acquisitions costs observably differ, once attributes x (including recovered variety quality) are accounted for. Thus, a variety that originates from China is handled and shipped by the same local warehouse worker as a variety produced in Alabama.

Equations (17) - (20) can be combined to solve for $p_{w,D}$, $p_{w,F}$, $c_{w,D}$ and $c_{w,F}$. This technique generalizes to the three high-level varieties used in the estimation.

[34] For details on markup calculations see Appendix D.

E Fixed Cost Details

I can compute this for every observed type in the data, however even with the limited data available, I may effectively only observe a few draws for ξ and ν. In particular, the locations of warehouses and different importing configurations makes a very large state space. As an alternative, I aggregate the state space to create ten options. For each wholesaler w in the data, I compute:

$$\pi_w(\mathbf{x}_w)\,|\mathcal{W} \quad \text{and} \quad \pi_w(\mathbf{x}_w)\,|\mathcal{W}' \tag{21}$$

where $\mathcal{W}$ is the set of wholesalers observed and $\mathcal{W}'$ is this set, plus an identical copy of the wholesaler w. I then aggregate and average each of these values across all firms to the ten aggregate observed wholesaler types in Table 8. The table displays estimated sets of upper and lower bounds and are not confidence intervals.

These bounds are empirically implemented by simulating counterfactual net variable profits $\pi_{\mathbf{x}}$ for each wholesaler configuration $\mathbf{x}$. This estimation technique can hypothetically provide extremely wide bounds. In practice, due to the number of wholesalers typically available in a market, bounds are relatively narrow, with the exception of the very largest wholesalers.[35]

Table 8 considers the lower and upper bounds of fixed entry costs $E_{\mathbf{x}}$ for various wholesaler configurations $\mathbf{x}$. While the underlying calculations are done by wholesaler market and industry, displayed results are averaged across markets. These results are further binned by broad groupings $\mathbf{x}'$. For clarity, wholesalers that only participate in international trade are combined with wholesalers that participate in both domestic and international trade.[36]

F Factory-less good manufacturers

Recent research (Bernard and Fort, 2015; Bernard et al., 2017) and anecdotal evidence suggest that the rise in wholesalers may be due to an economy-wide trend in former manufacturing firms closing domestic production operations and only retaining design and distribution facilities. It appears the trends captured in this paper are largely independent and highly complementary to the findings in Bernard and Fort (2015); Bernard et al. (2017). I address this research in three different ways. First, the residual quality term ξ may capture a portion of this change. Second, a large proportion of these former manufacturing firms are removed in the raw data. Third, the evidence from international sourcing patterns is inconsistent with common formulations of this outsourcing theory.

In the demand analysis the residual term ξ_w captures the quality of a wholesaler w that rationalizes its price and market shares. If these wholesalers use contract manufacturing and these

[35] Bounds can be computed for every every possible observed configuration of a wholesaler. However, as there are 2^{51} possibilities for wholesaler location choices, not all possible configurations are seen in the data. Selection of firms into 'positive' cells is a very real and possible problem. Thus I bin cells and average across the observed number of firms. Counterfactuals will only consider aggregate bins with positive firm counts.

[36] This binning does introduce potential compositional issues within each bin. Over time, the types of firms do change; firms in the biggest bin are on average 'larger' over time and firms in the domestic and international bins carry more product lines. This is reflected in the higher estimated variable profits and thus estimated entry costs.

contract manufacturers produce products with higher qualities, then the trend towards factory-less good manufacturing is captured in this analysis. This is plausibly one of the underlying mechanisms that deserves further study. However, it is not clear that these firms dominate the data.

The Census of Wholesalers includes categorizations such as "own-brand marketer" and "single-brand marketer". If these wholesalers market only their own brand, then they are excluded from the sample of wholesalers and treated as manufacturers. A possible example could be the electronic firm Apple, that markets its own products but outsources manufacturing.[37] In addition, the analysis also excludes manufacturer owned sales and branch offices. These locations exist to distribute products manufactured by a parent or sister firm. The elimination of these establishments does reduce the observed growth in the wholesale sector, providing a conservative approach to measuring the wholesaler market shares gains.

The behavior of the growth of these wholesalers takes a very particular form. As shown in tables A2 and A3, the largest wholesalers are importing many more varieties from new foreign sources and simultaneously increasing their distribution network within the United States. A common formulation of the factory-less good manufacturer theory is that these manufacturers close down production in the United States and move manufacturing abroad, with little to say about designing new varieties for production or expanding local distribution networks. As the benefit from wholesaling primarily derives from both sourcing new international varieties, not just moving production overseas, and expanding domestic distribution, it is unclear that the shift to factory-less production is driving the entirety of the trend towards wholesaling.

Finally, while this trend may be new for some firms, with Apple closing manufacturing lines in the United States and outsourcing manufacturing to Foxconn in China, such 'factory-less' producers have existed for a long time. Historically, when IBM produced personal computers, they did not produce all components sold with the IBM brand; the printer was simply a rebadged Epson device imported from Asia.[38]

G Endogenous Market Size

In the main model, the number of buyers of type j: $B_j \equiv B \times b_j$ is exogenous. This section endogenizes this aspect, to better line up with the macroeconomic and trade literatures.

Generally, discrete choice models assume that the total mass of possible purchasers remains constant. However, this assumption may not be plausible across all intermediate good markets. If a set of new wholesalers, perhaps supplying goods from a new foreign market enter, one could expect an increase in the overall downstream market size. I consider the elasticity of a market size for a customer j with respect to the valuation of all wholesaler options. While adopting a slightly different functional form, this stage follows Hausman et al. (1995), where consumers first choose quantity before choosing among a set of discrete choices. The quantity choice incorporates

[37] The exact categorizations of firms cannot be disclosed outside of the U.S. Census Bureau, it is unclear where firms such as Apple stand and the textual discussion is purely hypothetical.

[38] The IBM 5152 printer was a version of the Epson MX-80 printer

information from the choice set in a parsimonious manner and models a situation where customers must pick their purchase quantities before receiving their idiosyncratic cost draws ϵ.[39]

The number of purchases of type j varies with the set of available wholesalers $\mathbf{x}$. This allows for an increase in the number of purchases following increases in aggregate wholesale supplier quality.

$$b_j \quad = \quad B(\mathbf{x})$$

This relationship is parameterized by:

$$B_j = A_j \left(\underbrace{\sum_{g \in T} (D_g)^{1-\sigma}}_{D_W} \right)^{\phi} \tag{22}$$

Where B_j is the number of purchasers of type j, $D_{w,j}$ is the aggregate inclusive valuation of sourcing from a wholesaler of type t for a customer of type j relative to directly buying from a manufacturer, and ϕ is the elasticity of the number of purchasers relative to the aggregate valuation of purchases. In particular, as shown earlier, this form of two stage decision making is consistent with simple forms of cost minimization. As I only vary the quality and quantity of wholesalers, I normalize the valuation of buying from a manufacturer to 1. Denoting the term within brackets as D_w and taking logs:[40]

$$\log B_j \quad = \quad \phi \log [D_W] + A_j. \tag{23}$$

The discrete choice setup allows us to directly estimate D_W using the market share of direct manufacturer shipments:

$$s_{0|j} = (D_W)^{-1}.$$

Thus I obtain the relationship:

$$\log B_j = -\phi \log \left[s_{0|j} \right] + A_j. \tag{24}$$

G.1 Estimating Market Size

I seek to (a) estimate the elasticity ϕ of the number of downstream purchasers with respect to the aggregate mean utility from wholesalers and (b) recover the the size of the market without wholesalers, $\mathbf{A}$.

Estimation uses equation (24), reproduced below:

$$\log B_j = -\phi \log \left[1 - S_j^W \right] + \log [A_j].$$

[39] In Hausman et al. (1995), vacationers choose the number of trips to take, which follows a poisson process that uses the inclusive values D from an earlier stage.

[40] This functional form is useful in that $\delta_{w,j}$ is only defined up to an additive constant. Since D_w is a summation of $\exp(\delta_{w,j})$, $(D_W)^{\phi}$ is defined up to a multiplicative constant.

This equation shows that the relative value of wholesalers compared to direct sourcing is entirely captured by aggregate wholesaler market shares.[41] The object of the estimation is to provide A_j for use as an instrument in the discrete choice estimation and parameter ϕ to identify the elasticity of aggregate demand. To better explain the identification strategy, I first elaborate on the level of observation. Each j is composed of three elements: downstream product category c (which is defined at the year-product level), downstream location l, and downstream purchase quantity q. Denoting $B_{c,q,l}$ as the total observed downstream purchases and $S^W_{c,q,l}$ as the aggregate wholesaler purchase share for product c, in region l, where the shipment size is q units, I estimate the following relationship:

$$\log B_{c,q,l} \quad = \quad -\phi \log\left[1 - S^W_{c,q,l}\right] + \lambda_{c,l} + \lambda_{c,q} + \lambda_{l,q} + \lambda_{c,q,l}. \tag{25}$$

The covariate $\lambda_{c,l}$ represents a fixed effect for a particular product c sold in region l, $\lambda_{c,q}$ represents a fixed effect for a particular product c sold at quantity q, and $\lambda_{l,q}$ represents a fixed effect for shipments of quantity q in a given region l. These covariates represent the local demand for certain products, the general nature of that demand, and the market size of that downstream location. The last term $\lambda_{c,q,l}$ represents the **deviation** of a particular (c, q, l) from the three previous fixed effects. The residual term A_j equals $\exp(\lambda_{c,l} + \lambda_{c,q} + \lambda_{l,q} + \lambda_{c,q,l})$, where the first three linear terms are controlled for, but the last term is unobserved. I then collect the set of residual demand shifters in vector $\mathbf{A} = \{A_j\}$.

Estimation assumes that $E[X_D \lambda_{c,q,l}] = 0$, where X_D includes share of goods sourced from wholesalers and the three fixed effects. Econometrically, the last lambda, $\lambda_{c,q,l}$ is not controlled for and may be correlated with wholesaler market shares. A related econometric risk is reverse causation: higher demand B may induce more wholesaler entry. Due to the timing assumptions made, structure of demand and explicit product-location fixed effects controlling for wholesaler and overall downstream demand presence, I explicitly rule this out. An alternative view of this assumption is that aggregate demand shocks affect both large and small purchases similarly; the difference between large and small purchases is entirely accounted for by wholesalers.[42]

G.2 Market Size Results

Estimates for the elasticity of the downstream market size with respect to expected utility from wholesaling are reported in Table A8. Columns (1) - (4) report results across various specifications.

[41] The expected utility in such discrete choice models is simply the inverse market share of the choices: $EU_j = 1/\left(1 - S^W_J\right)$.

[42] Identifying variation can be summarized as follows. Consider the sales of industrial chemicals in Connecticut. Estimation looks at the deviation in the number of large and small orders from both the Connecticut averages for those orders, as well as at the deviation within industrial chemicals. Additionally, in contrast to the sixty product markets (over three years) used in the discrete choice estimation, a more refined set of over 400 products are used in this estimation.

An alternative instrumentation strategy would be to use geographic variables exploiting changes in wholesaler costs across regions, as done in the last demand stage. For robustness, data is aggregated up to the product-location level and the suggested instrumentation strategy is used, dropping product-location fixed effects. While the magnitude of ϕ is slightly larger, results are broadly similar.

Table A8: **Market Size Estimation (1st Demand Stage)**

	1	2	3	4
Elasticity $-\ \phi$	0.234	0.174	0.248	0.262
	(0.020)	(0.045)	(0.017)	(0.029)
Weighted	N	Y	N	Y
Aggregation Level	SCTG-4		SCTG-6	
Fixed Effects	Product-Year × Location			
	Product-Year × Shipment Size			
	Location × Shipment Size			

Notes: Regression results use the logarithm of total market size as the dependent variable. Robust standard errors clustered by product-year. See text for full regression specification. Aggregation by Standard Classification of Transported Goods (SCTG).

Shipments are binned in the same nine size categories as in the demand choice estimates. Locations consider the fifty U.S. states as well as the District of Columbia. Product-year categories consider Standard Classification of Transported Goods (SCTG) good classifications, which are more disaggregated than the wholesaler NAICS categories used in the demand choice estimation. Columns (1) and (2) consider 4-digit SCTG categories, while columns (3) and (4) consider 5-digit SCTG classifications. In general, more disaggregated classifications lead to more fixed effects and higher R^2 values, even though the parameter estimates do not significantly change. Columns (2) and (4) weight results based on market size.

In general, all four specifications find precise parameter estimates for the elasticity ϕ between .25 and .30. If implemented in the main estimation, these estimates imply about 20% higher welfare gains - within the same order of magnitude.

H Endogenous Quality

In the main model, quality deviations ξ are exogenous. I propose a mechanism whereby ξ is endogenously chosen by firms. Suppose between Stage 1 and Stage 2, firms choose ξ. Call this Stage 1.5. While theoretically easy to add, this stage presents estimation challenges and requires a modified estimation technique. In particular, this restricts the parameters estimated in the demand estimation stage. Instead of finding valuations for firm attributes $x_{w,i}$, all attributes are subsumed in a single vertical quality dimension ξ. Therefore now:

$$\delta_{w,i} = \alpha p_{w,i} + \xi_{w,i}.$$

H.1 Model Changes

Now, firms choose market entry in two stages. First, wholesalers choose their domestic distribution locations entering as a firm with domestic sources, international sources, or with both domestic and international sources. In the second stage, firms choose the quality of their products, and their internationally and domestically sourced varieties. This includes the variety of products a wholesaler offers as well as the consumer service provided by the wholesaler. In terms of the model, a firm must optimally choose $\xi_{w,i}$ for both their domestically and globally sourced products.

Conditioning on a firm's type and location choices, the model assumes wholesaler w optimally chooses $\xi_{w,i}$ for each product i. In particular they must invest $f_w(\xi)$ to receive product attributes $\xi_{w,i}$, which realize in operating profits $\pi_w(\xi_{w,i})$. If a firm only participates in domestic sourcing, they maximize the following problem by choosing their optimal firm quality $\xi_{w,i}$:

$$\max_{\xi=[\xi_{w,D},0,0]} \pi_w(\xi) - f_w(\xi) \tag{26}$$

If firms participate in both first-world global and domestic markets, a firm w must choose two parameters, $\xi_{w,n}$ for $n \in \{F_H, D\}$, where $n = F_H$ represents first-world imports and $n = D$ represents domestically sourced products:

$$\max_{\xi=[\xi_{w,D},\xi_{w,F_H}]} \pi_w(\xi) - f_w(\xi) \tag{27}$$

For simplicity, I now present results involving a single firm only involved in domestic sourcing and suppress firm subscript w and product type subscript i. Conditional on location choices (market entry), a firm's profit maximization produces first order conditions:

$$\frac{d\pi(\xi)}{d\xi} = \frac{df(\xi)}{d\xi} \tag{28}$$

Without any errors, this solution concept implies that any two ex-ante identical firms will choose the same ξ. As firms are only differentiated on an extremely limited set of dimensions in the market entry stage, this setup will not fully rationalize the data. To better rationalize the data and account for the heterogeneity present in the world, the model allows for firm-specific investment cost shocks. Before wholesalers choose their market position, but after entering the market, each wholesaler receives shocks to the marginal costs of investing. Call these shocks η_ξ.

Given these shocks, two ex-ante firms will no longer make the same investment choices and thus fully rationalize the observed data. Given a form for a time-varying investment function $f(\cdot)$, parameterized by the vector χ, the econometrician can recover changes in the return to investment. In particular, in the context of wholesaling, are the returns to investing in domestic and international quality differentially changing for large and small firms?

H.2 Estimation

Unobserved downstream consumer valuations ξ are not exogenous shocks as in standard discrete choice models. They are the product of wholesale firm investments. This ξ is better written as $\xi(\mathbf{x})$. In this case, all fixed effects and ξ are all subsumed by the new measure $\xi(\mathbf{x})$. $\xi(\mathbf{x})$ is no longer a residual, it is a complete measure of quality. Regardless, the coefficient α can be identified as a cost shock hits a particular firm following their choice of $\mathbf{x}$ and $\xi(\mathbf{x})$. In terms of β_q, β_l, and σ, they are identified from aggregate moments. As α is the only coefficient required to derive demand elasticities, estimation can proceed in a more restricted fashion.

Having made these assumptions, identification of this investment function proceeds directly from the first order conditions in equation (28). For any given company configuration $\mathbf{x}$, assume that the fixed costs of market positioning are:

$$f_w^x(\xi, \eta) = \left(\frac{\chi_1^x}{\chi_2^x}\eta_{w,\xi}\right)\exp\left(\chi_2^x \xi\right) + E_a$$

The function $f_w^x(\xi)$ measures the cost of investing in quality ξ for wholesaler of configuration x. There are scalar fixed costs E_x and two parameters, χ_1^x and χ_2^x. Finally there is a wholesaler specific shock $\eta_{w,\xi}$. This structural investment cost shock is known to the firm, but not the econometrician.

Conditional on entry, a wholesale firm of configuration x seeks to maximize profits $\pi_w(\xi)$ net of investment $f_w^x(\cdot)$. As both $\pi_w(\cdot)$ and $f_w^x(\cdot)$ are smooth linear functions, computation of the optimal profits requires solving the firm's first order conditions. Marginal investment costs are:

$$\frac{df_w^x(\xi, \eta)}{d\xi} = \left(\chi_1^x \eta_{w,\xi}\right)\exp\left(\chi_2^x \xi\right)$$

and marginal profits stem from the first derivative of equation (5) with respect to ξ, $d\pi_w(\xi)/d\xi$. As all the parameters in $\pi(\cdot)$ are known, the optimal marketing costs in equilibrium solve:

$$\frac{d\pi_w(\xi)}{d\xi} = \frac{df_w^x(\xi, \eta)}{d\xi} = \left(\chi_1^x \eta_{w,\xi}\right)\exp\left(\chi_2^x \xi\right). \tag{29}$$

Taking the logarithm of this equation produces the following relationship:

$$\log \frac{d\pi_w(\xi)}{d\xi} = \log \chi_1 + \chi_2 \xi + \log \eta_{w,\xi}. \tag{30}$$

The relationship should be theoretically estimated by ordinary least squares, however the shock $\eta_{w,\xi}$ likely is correlated with the choice of ξ. This echoes the endogeneity problem with ξ and h_w in estimating equation (9). Estimation of χ requires a shifter of ξ that is uncorrelated with η. This leads to an assumption required for identification.

Assumption 4 *There exist Z_η such that $E[\eta Z_\eta] = 0$.*

Thus, under this model's demand and supply systems, investment cost parameters χ are identified.

What is a plausible exogenous shifter of ξ? Estimation could use a combination of two shifters, one using the timing of the game and the second using geographic differentiation. The first shifter is similar to the cost shifters in the demand estimation. Wholesale firms are likely to choose higher levels of ξ when similar wholesale firms in nearby, but unrelated markets choose higher levels of ξ. So the average ξ in New Haven for importing chemical wholesalers can be used as an instrument for New Haven electronic wholesalers. The second shifter exploits the timing of the game. Firms choose their attributes $\mathbf{x}$ before investing in ξ, thus the number of firms of type $\mathbf{x}'$ at the state, regional, and national level shift the choice of ξ independently of η.

In computation, $\pi_w(\xi)$ is not fully known by a firm before the investment decision ξ is made. There is an unobserved cost shock ν from equation (10) that shifts profits. I assume the distribution of ν is known and firms maximize their expected profit. To aid in computation, instead of numerically integrating over ν, simulated draws of ν are used to compute $E[\pi_w(\xi)]$. For simplicity, I omit the expectation in what follows.

Investment function $f_w^x(\cdot)$ is identified up to some fixed entry constant E_x. Following estimation of χ_1^x and χ_2^x , this step generates the distributions $G_\eta^x(\cdot)$ for investment shocks of $\eta_{w,\xi}$. I denote ξ_w^* as the optimal choice for firm w with investment cost shocks η.[43]

Second-stage net profits for a firm of configuration c are

$$n_a(\eta) = \pi_w^x(\xi^*(\eta)) - \bar{f}_w^x(\xi^*(\eta), \eta),$$

where $\bar{f}_w^x(\cdot) = f_w^a(\cdot) - E_c$.

Note that $f(\cdot)$ is only identified up to some constant E_x, $\bar{f}(\cdot)$ subtracts this constant. The function $n_x(\eta)$ is used in the next stage to identify this entry cost E_x. For tractability, I assume that fixed cost E_x is not paid in this stage, as firms in this stage have already entered into the market and that an infinitesimally small investment in ξ (that is $\xi \to -\infty$) will realize a investment cost of 0.[44]

[43]The chosen functional form for $f_w^a(\cdot)$ and the estimation equation (30) imply that $\chi_1\eta$ is greater than zero, thus as long at χ_2 is greater than zero, $f_w^a(\xi^*)$ will be always greater than zero.

[44]Additionally, under a free entry condition for counterfactuals, estimates from this step are not needed to compute alternative equilibria. Due to free entry, firms will reenter until $\pi'(\xi) = F'(\xi)$. This step does matter for when the fixed costs of entry change, but market positioning costs are unaltered. This step is mostly critical for understanding the role of 'business' stealing arising from competition.