Triadic Gravity in Intermediated Petrochemical Trade*

Hyungjin Kim† Yang Shen‡

April 2026

Abstract

How does intermediated trade alter the response of bilateral trade flows to cost shocks? We show that when firms can reroute shipments through intermediary hubs, the bilateral trade elasticity is strictly attenuated relative to the standard gravity benchmark—attenuated by a factor equal to the share of trade currently flowing directly between origin and destination. We establish this result theoretically and validate it using transaction-level data from a Korean petrochemical trading house with 25 offices worldwide, where we observe the complete buyer–seller–intermediary triad for every shipment. In this network, 84.5% of origin–destination pairs route all trade through hubs, so a direct-route cost shock leaves their bilateral trade volume entirely unchanged—while standard gravity would predict a large negative response for those same pairs. Aggregating across all routes, standard gravity overstates the bilateral trade-flow response to a uniform cost shock by roughly an order of magnitude in this network (value-weighted factor of 6.5; range 3×–13× depending on the substitution elasticity). These findings imply that bilateral gravity models systematically overstate trade losses from tariff increases in sectors with deep intermediation.

JEL Codes: F10, F23, L81

Keywords: Intermediated trade; gravity equation; firm-to-firm transactions; trade elasticity

*We thank the participants at the Korea International Economic Association Winter Conference, the International Economics and Finance Society (China), and the Asia-Pacific Symposium on Economics and Finance for helpful comments and suggestions. All errors are our own.

†Korea Development Institute. Email: hyungjinkim@kdi.re.kr.

‡College of Economics and Finance, Hanyang University. Email: shenyang@hanyang.ac.kr.

1 Introduction

Consider a shipment of polyethylene from a Saudi refinery to a Vietnamese plastics firm, settled through a Korean trading house's Singapore office. Standard gravity records a single bilateral distance—Riyadh to Ho Chi Minh City—while the actual freight path traverses two legs: Riyadh to Singapore, then Singapore to Ho Chi Minh City. When bilateral compliance and logistics costs on the direct Saudi–Vietnam route rise, the trading house does not reduce shipments in proportion; it shifts transactions to cheaper hub routes, absorbing the shock through rerouting rather than reducing trade. The bilateral trade elasticity is *attenuated*—potentially to zero—by the existence of indirect routing alternatives. Intermediaries mediate 8 to 20 percent of exports in advanced economies and substantially more in emerging markets,[1] yet their geography is invisible to standard bilateral gravity.

This paper establishes that the attenuation is empirically first-order and theoretically tractable. The bilateral trade elasticity is bounded strictly between zero and the standard single-channel benchmark, with the degree of attenuation governed by the direct-routing share—the fraction of trade between a given origin and destination that currently flows without a hub. When all trade flows directly, the standard gravity elasticity is fully recovered. However, when all trade flows through hubs, bilateral cost shocks do not reduce trade: firms reroute rather than exit. Standard gravity, which implicitly treats all trade as direct, therefore overstates trade-policy responses in proportion to the degree of indirect routing.

Using a unique transaction-level dataset from a Korean trading house with offices in 25 countries, covering monthly records for 23 petrochemical products between 2018 and 2022, we observe the complete triad—buyer, seller, and intermediary—behind every shipment. Petrochemical intermediation is a transparent laboratory for the routing-mode mechanism: products are commodity-like with well-documented elasticity ranges, hub geography is concentrated (Korea and Singapore dominate the network), and the firm's complete routing decisions are observed rather than inferred from customs aggregates. The data reveal that buyers, sellers, and intermediaries are often located in different countries, generating three distinct pairwise distances rather than one. Leveraging this structure, we extend the canonical gravity framework to a triadic gravity equation that explicitly

[1] See Bernard et al. (2010), Ganapati (2025), Bernard et al. (2015), and Crozet et al. (2013).

embeds origin–destination, origin–intermediary, and intermediary–destination distances. Because all transactions flow through a single trading house, quantitative magnitudes are specific to this firm's network; the mechanism and the attenuation bound are expected to generalize.

Our estimation reveals that the intermediary's location fundamentally reshapes the operation of gravity. While physical distance and cultural proximity remain influential, their effects are highly asymmetric across the three legs. We find that intermediaries selectively bridge the longest geographic gaps: once the hub's location is accounted for, the penalizing effect of the direct exporter–importer distance shrinks or even reverses. Trade volumes increase with the distance between exporters and their hub—consistent with hubs pooling supply from geographically remote sellers—while markups rise with the hub's distance to the buyer, consistent with longer last-mile legs commanding higher service fees.

To interpret these patterns structurally, we develop a model of hierarchical mode selection. Firms choose between direct exporting and indirect routing based on a trade-off between variable transport costs and fixed access fees charged by intermediary offices. The central mechanism is a switching margin: when direct trade costs rise, marginal firms shift to the indirect route rather than exit the market, thereby dampening the bilateral trade elasticity relative to the standard model.

This switching mechanism has a precise structural implication: the bilateral trade elasticity is proportional to the direct-routing share, ranging from the full standard-gravity value (when all trade is direct) down to zero (when all trade flows through hubs). We show that the key trade-elasticity parameter is identifiable even when the data record only the chosen routing mode, by constructing a *Switch* control—the cost advantage of the direct route over the cheapest hub alternative, built from CEPII great-circle distances—that absorbs the mode-selection effect and allows consistent estimation from the direct-trade subsample.

These findings refine the interpretation of gravity coefficients and carry direct policy implications. Standard models calibrated on bilateral elasticities effectively assume that high trade costs lead to market exit. Our triadic framework reveals that trade is more resilient: bilateral shocks propagate through the network, attenuated by the scope for rerouting, with the attenuation strongest for pairs that currently trade mostly indirectly. Suggestively, infrastructure investments that lower compliance costs at third-country hubs could alleviate geographic barriers for spokes that lack the scale to export directly. As an illustrative episode, we document hub reorganization

following Russia's sharp decline as a petrochemical intermediary after 2021. This descriptive event study highlights which hubs absorbed displaced volume and provides evidence consistent with the model's predictions about rerouting, although we do not estimate a structural counterfactual for this episode.

Related literature. Our study contributes to three strands of research. First, it speaks to the classic gravity literature that rationalizes bilateral trade with bilateral trade costs (Anderson and van Wincoop, 2003; Eaton and Kortum, 2002). While a large body of work estimates gravity with firm-level data, most studies collapse intermediary activity into exporter fixed effects or dummies (Bernard et al., 2010, 2015; Crozet et al., 2013). The mode-selection aggregation in our model is related to the extensive-margin aggregation in Chaney (2008), where the firm-size distribution shapes the aggregate distance elasticity; we extend this logic to multi-mode routing and show that the presence of an indirect option attenuates rather than amplifies the bilateral elasticity. By estimating a "triadic-gravity" equation, we show that ignoring the intermediary's location biases distance elasticities and misattributes the role of geography in trade costs.

Second, we add to the literature on intermediated trade. Theoretical work highlights three key functions of intermediaries: reducing fixed entry costs (Ahn et al., 2011; Irarázabal et al., 2013), exploiting economies of scope (Akerman, 2018), and overcoming informational frictions (Rauch, 1999; Felbermayr and Jung, 2011; Antras and Costinot, 2011; Antras and Chor, 2013). We complement these findings by focusing on a distinct dimension: the *location* of intermediary hubs and how hub geography shapes trade flows, independently of the matching, bargaining, or information-exchange functions that prior work emphasizes. This perspective reveals that the hub's position—its distance to the origin country and to the destination country—independently shapes both the volume and price of trade in ways that standard bilateral gravity cannot capture.

Third, we connect to the growing literature on network and multi-leg trade costs (Bernard et al., 2019; Antras and Chor, 2013; Fally and Hillberry, 2018). Our firm-level anatomy of how trade splits across modes and destinations relates to Eaton et al. (2011), who decompose the participation margin of French exporters across markets; our contribution is to add a routing dimension to that anatomy by tracking the intermediary leg explicitly. The Russia hub-substitution episode in Section 5 also connects to the literature on how trade-cost shocks reshape supply chains (Con-

coni et al., 2018): as sanctions altered compliance incentives, routes were reorganized across hub jurisdictions in a pattern consistent with our model's predictions about rerouting over exit. To our knowledge, Blum et al. (2010) is the only prior paper with firm-level observation of all three nodes in a trade triad simultaneously; they use Chilean customs data to document that intermediated trade is geographically concentrated and strongly size-sorted. We extend their framework in three dimensions: our panel structure allows us to track buyer–seller–intermediary triads over time; we observe the intermediary's realized profit margin, enabling markup regressions that their data do not support; and we derive and estimate a structural sufficient statistic that recovers the Pareto shape parameter from the selection margin, which goes beyond their descriptive evidence.

The remainder of the paper proceeds as follows. Section 2 describes the data and empirical patterns that motivate our specification. Section 3 develops the dyadic and triadic gravity models and presents Poisson pseudo-maximum-likelihood (PPML) estimates. Section 4 presents the theoretical framework and derives the monotone sorting condition and the attenuated trade elasticity. Section 5 presents two quantitative exercises: one quantifying the attenuation of bilateral cost shocks relative to the standard gravity benchmark, and one using the structural hub-demand formula to back out the implied compliance-cost increase from Russia's observed decline as an intermediary after 2021. Section 6 concludes.

2 Data and Empirical Patterns

2.1 Data

Source and coverage. We use the internal transaction ledger of a major Korean intermediary[2] specializing in petrochemical trade. The raw data records every shipment handled through the firm's 25 international offices between January 2018 and December 2022.[3] Each row represents a contract–month observation and contains (i) the seller firm i, (ii) the buyer firm j, (iii) the product g at the HS six-digit level, (iv) the intermediary office m that facilitated payment and logistics,

[2]Throughout the paper we use the terms *intermediary*, *trading house*, and *wholesaler* interchangeably to refer to this firm and its role; *intermediary* is preferred in the theoretical sections, *trading house* when referring specifically to the data source, and *wholesaler* when connecting to the broader literature.

[3]These offices are located in Australia, Brazil, Chile, China, Germany, Ecuador, Indonesia, India, Ireland, Iran, Japan, Kazakhstan, Cambodia, South Korea, Mexico, Philippines, Poland, Russia, Singapore, Turkmenistan, Turkey, Ukraine, United States, Vietnam, and South Africa.

and (v) shipment-level quantities, invoice values, sales, and profits earned by the intermediary. Unique firm identifiers allow us to track repeated dyads or triads over time. After removing test transactions, returns, and within-month duplicates, the cleaned panel comprises approximately 14,000 shipments with a total value of USD 3.57 billion. Table 1 summarises the dataset.

Table 1: Dataset overview

Coverage	
Sample period	2018–2022
Petrochemical products	23
Intermediary offices	25
Transaction volume (cleaned panel)	
Shipments (contract–months)	13,776
Total trade value (USD bn)	3.57

Notes: Each observation is a contract–month pair linking a seller firm, a buyer firm, and an intermediary office. Shipment count and total value are after removing test transactions, returns, and within-month duplicates. Product-level breakdowns are in Table A1.

Triadic geographic variables. Let o, d, and m denote the countries of the seller, buyer, and intermediary office, respectively. For each transaction linking seller i (in origin country o), buyer j (in destination country d), and intermediary m, we compute three great-circle distances–seller-buyer (distance_{od}), seller–intermediary (distance_{om}), and intermediary–buyer (distance_{md})–from CEPII's `geo_dist` dataset (Mayer and Zignago, 2011).[4] We further attach bilateral characteristics that may differ across the three legs: common language, colonial ties, and shared borders, all defined at the country level and assigned to each firm-to-firm shipment according to its origin-intermediary-destination structure.

2.2 Product-level empirical patterns

We begin by presenting descriptive statistics at the product level and empirical patterns that reveal substantial heterogeneity in market size and trading frictions across products.

[4]For sellers and buyers, we distinguish firm and country indices. For intermediaries, however, we use the same index m for both the firm and its country of residence, as the data only report the location of the intermediary office that conducts the petrochemical shipment.

Transaction size and network breadth. The firm's ledger covers 23 distinct petrochemical products spanning fuels, polymers, and fertilizers. Average trade volumes, sales values, and intermediary profits vary substantially across products, as does the breadth of the buyer and seller network. Full descriptive statistics by product are reported in Table A1 in the appendix.[5]

Three patterns emerge. First, the intensive margins of intermediary trade, measured by quantity and value, vary substantially across products, reflecting heterogeneity in market size and trading conditions. Second, the breadth of the intermediary's network also differs widely. Products such as butyl rubber, jet oil, and motor gasoline (MOGAS) involve only a few trading partners, while polyethylene (PE) and domestic bunkering link dozens of distinct buyers and sellers. In most cases, the number of buyers exceeds the number of sellers. Third, the intensive margins do not rise in proportion to the breadth of firm-to-firm transactions. Some products, such as olefins and complex fertilizers, exhibit large trade volumes and values despite involving relatively few buyers and sellers, suggesting high concentration or repeated exchanges within stable trading relationships. Bunkering abroad reports a negative average intermediary profit (-5.17); this reflects the accounting convention in which the trading house acts as a logistics coordinator rather than a price-setting intermediary for some bunkering contracts, recording net service fees that turn negative when fuel procurement costs exceed the billed service charge in a given transaction.[6]

Realized buyer–seller matches across products. Intermediaries facilitate matches between buyers and sellers, but the fraction of potential relationships that are actually formed varies sharply across products. As shown in Figure 1, for most products fewer than ten percent of possible matches are ever activated, which is consistent with sizable fixed or search costs of establishing and maintaining relationships. In contrast, olefin and tarpaulin reach match shares of around twenty percent, suggesting relatively standardized specifications and lower search frictions. Plastic-related products (e.g., PE, PP, PET, and plasticizers) remain far more selective, with only a small fraction of feasible matches realized, in line with heterogeneous grades, certification requirements, or more volatile, relationship-specific demand.

[5]The firm's transaction ledger also records four additional item groups—fine chemicals, potassium fertilizer, tires, and biofuel—which we exclude from the analysis. Fine chemicals and biofuels are peripheral to the firm's core petrochemical mandate; potassium fertilizer averages fewer than five transactions per year; tires lack sufficient counterparty coverage to reliably identify buyer and seller countries. Including them does not materially affect any of the results.

[6]In Section Section 3, negative-profit bunkering observations are dropped from the markup regression (as markup $<$

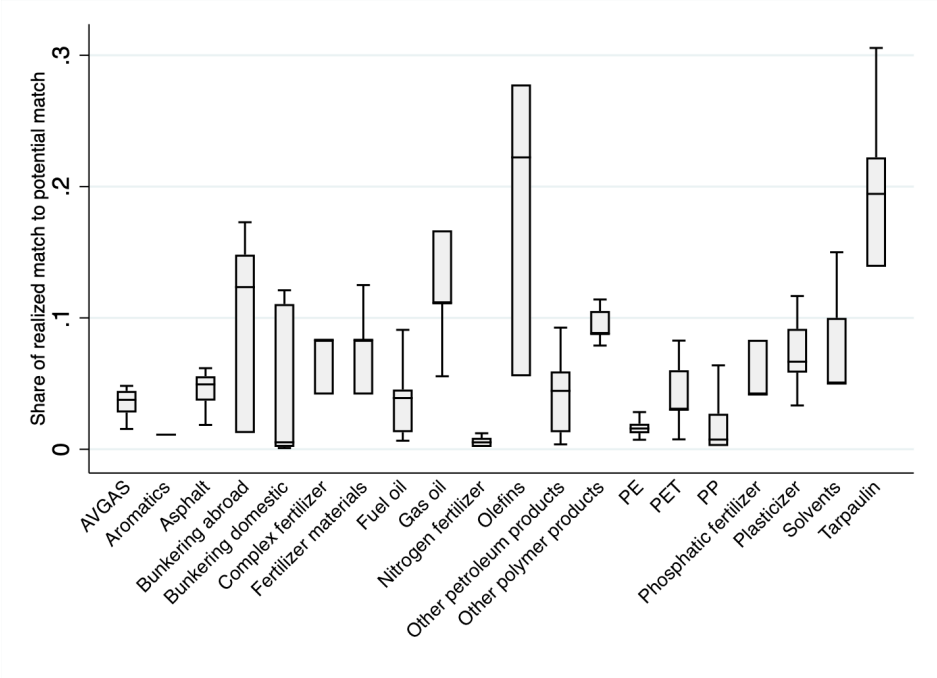


Figure 1: Share of realized buyer–seller matches among potential pairs, by product. For each product-month, the match share is the number of active buyer–seller pairs divided by the total number of potential pairs.

These product-level patterns suggest that the intermediary's margin and matching intensity depend on market thickness and the degree of product standardization. In the empirical analysis that follows, we account for this heterogeneity by including product fixed effects.

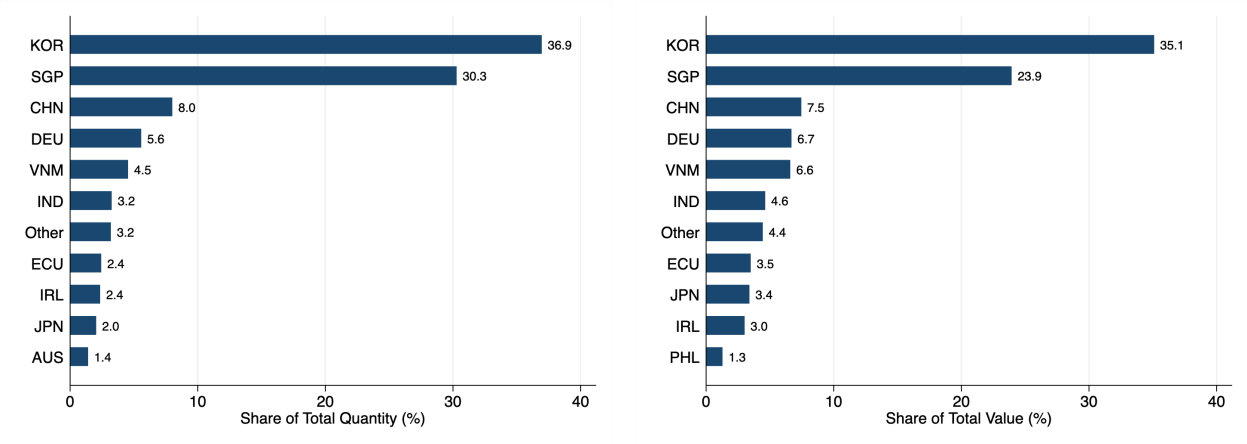
2.3 Country-level Empirical Patterns

We now document empirical patterns at the country level, illustrating how intermediaries link sellers and buyers across borders and how these relationships vary with distance, sales, and pricing outcomes.

Geographic concentration of intermediary trade. Intermediated petrochemical trade is highly geographically concentrated. Figure 2 shows the distribution of total transaction value and physical quantity by the location of the intermediary office. South Korea and Singapore together account for more than half of all activity, with China, Germany, and Vietnam forming a second tier of intermediary hubs. The pattern is nearly identical whether measured in metric tons or in

0 observations are excluded from all PPML specifications) but are retained in the quantity and value regressions.

U.S. dollars, indicating that the dominance of these countries reflects the sheer scale of shipments routed through them rather than price differences, product specialization, or markup variation. These locations host large trading houses with established logistics networks, storage capacity, and financial clearing functions, making them natural aggregation and redistribution centers in global petrochemical supply chains.



(a) Intermediary location ranked by quantity

(b) Intermediary location ranked by value

Figure 2: Intermediated petrochemical trade by intermediary location

Cross-border fragmentation and co-location. Table 2 summarizes 13,776 intermediary transactions recorded between 2018 and 2022, classified by co-location patterns between the origin (o), destination (d), and intermediary (m). For each year, the first row reports the number of transactions and the second row their corresponding shares. We distinguish three partially fragmented cases: domestic trade between seller and buyer ($o = d$), buyer–intermediary co-location ($d = m$), and seller–intermediary co-location ($m = o$), alongside the fully fragmented case in which all three countries differ ($o \neq d$, $d \neq m$, $m \neq o$) and the fully domestic benchmark ($o = d = m$).

Three patterns stand out. First, intermediation has become increasingly destination-centric. Transactions with $d = m$ already accounted for about 40 percent of all matches in 2018. They rose to nearly 75 percent by 2022, suggesting that intermediaries increasingly embed themselves in final markets to handle customs, compliance, and after-sales services. Second, seller-side intermediation is declining. The share of $m = o \neq d$ falls from roughly 14 percent to 2 percent over the period, so exporter-side wholesalers account for only a small slice of trade. Third, fully domestic matches

Table 2: Annual patterns of intermediary transactions by co-location type

Year	Total	Fragmented ($o \neq d$, $d \neq m$, $m \neq o$)	Partially Fragmented ($o = d \neq m$)	($d = m \neq o$)	($m = o \neq d$)	Domestic ($o = d = m$)
2018	5100	1836	204	2016	732	312
		(0.360)	(0.040)	(0.395)	(0.144)	(0.061)
2019	2976	1008	72	1524	180	192
		(0.339)	(0.024)	(0.512)	(0.060)	(0.065)
2020	1956	708	84	924	84	156
		(0.362)	(0.043)	(0.472)	(0.043)	(0.080)
2021	1908	564	84	1236	0	24
		(0.296)	(0.044)	(0.648)	(0.000)	(0.013)
2022	1836	324	60	1368	36	48
		(0.176)	(0.033)	(0.745)	(0.020)	(0.026)
All	13776	4440	504	7068	1032	732
		(0.322)	(0.037)	(0.513)	(0.075)	(0.053)

Notes: Column "Total" reports the overall number of matched transactions. Symbols o, d, and m denote the origin, destination, and intermediary countries, respectively. Column "Fragmented" corresponds to transactions in which all three parties are located in different countries, while column "Domestic" corresponds to fully domestic transactions. Columns under "Partially Fragmented" indicate cases in which two parties share a country while the third is located elsewhere. Numbers in parentheses represent shares of the yearly total.

remain rare even within this Korean-owned network, underscoring how globalized petrochemical supply chains have become. Section A.2 documents similar patterns for traded quantities, value, and intermediary margins.

It is worth noting that Table 2 also shows a steady decline in the number of intermediated matches between 2018 and 2022. Two natural concerns arise when interpreting this pattern. One possibility is that the petrochemical sector experienced an aggregate contraction, leading to fewer transactions overall. Another possibility is that incumbent buyer-seller pairs may have reduced their reliance on intermediaries as they grew, gradually shifting toward direct trading. Although a decline in intermediated sales alone cannot, in principle, distinguish between aggregate contraction and channel substitution, two pieces of evidence suggest that neither industry-wide shocks nor expansion-driven graduation is the primary driver of exit from intermediated trade.

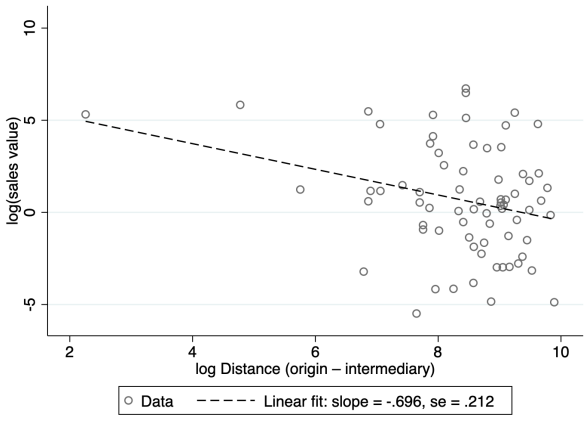
First, aggregate imports and exports of petrochemical products in the three major intermediary countries–Korea, Singapore, and China–show no systematic decline over this period (Figure A1). This rules out the possibility that the falling number of matches reflects a shrinking market.

Second, the firms that exit the intermediary network do not display the growth profile typically associated with "graduation" into direct trading. Both sellers and buyers who leave intermediary relationships show declining average monthly sales in the months leading up to exit (Figure A2), and exiters are systematically smaller than stayers in the 12 months preceding their final transaction (Figure A3). This stands in contrast to the well-documented pattern in which large or expanding firms are the ones most likely to bypass intermediaries.

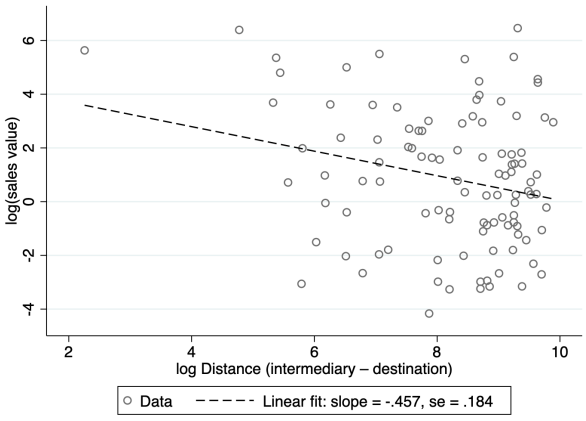
A third, more likely explanation for the post-2020 decline is geopolitical disruption in petrochemical trade routes. Russia served as an intermediary hub for this trading house beginning in 2019, reaching a peak of 72 transactions in 2020 (USD 6.6 million) and a value peak of USD 22.2 million in 2021, before declining sharply to 24 transactions (USD 5.0 million) by 2022—a pattern consistent with the impact of sanctions-related compliance risk on petrochemical routing through Russia-adjacent hubs. The decline in Russia's intermediary role, combined with compositional shifts toward China and Vietnam as replacement hubs, is the most plausible driver of the observed network contraction (see Section 5.2). Critically, this compositional change does not bias the triadic gravity estimates if the underlying distance elasticities are stable across the pre- and post-2020 periods. Section A.8 provides a formal test: adding interactions of each distance leg with a post-2020 indicator to the main triadic PPML regression, the three interaction terms are jointly insignificant at the 5% level ($p = 0.071$), supporting the interpretation that the gravity coefficients are stable across the network disruption period.

Distance and sales value. Figure 3 plots the relationship between country-pair distances and sales values (in millions of U.S. dollars) aggregated across firms and years from 2018 to 2022. Each panel corresponds to one leg of the triad: origin–intermediary, intermediary-destination, and origin–destination. In all three cases, logged sales values decline sharply with distance. The estimated coefficients, ranging from -0.46 to -0.76, imply that a 1% increase in great-circle distance reduces total sales by 0.46–0.76%. Equivalently, doubling distance reduces total sales by approximately 27–41% (since $2^{\beta} \approx 0.59$–0.73).

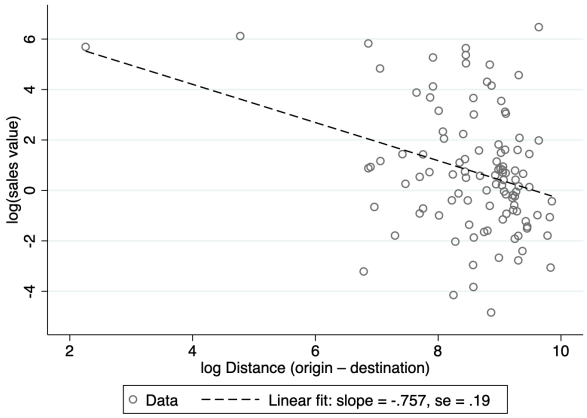
This negative relationship is remarkably consistent across all three legs, suggesting that geographic separation limits not only direct trade between buyers and sellers but also the efficiency of intermediation. Similar patterns emerge when using alternative outcomes, including the number



(a) Distance between origin and intermediary

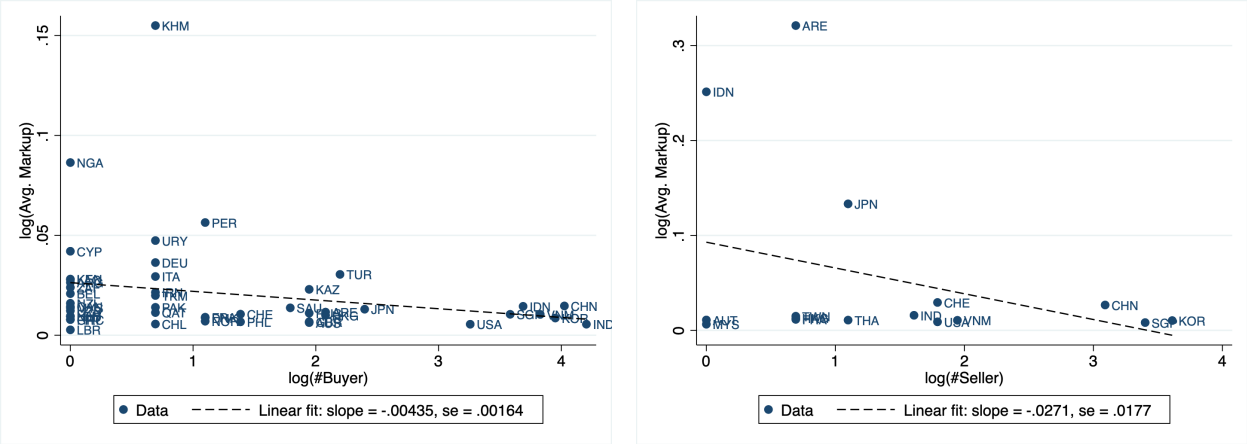


(b) Distance between intermediary and destination



(c) Distance between origin and destination

Figure 3: Logged triadic distance and logged sales values



(a) Markup per buyer

(b) Markup per seller

Figure 4: Average intermediary markup per buyer and per seller

of transactions, traded quantities, and intermediary profits (see Section A.4). Together, these patterns underscore the importance of a triadic view of trade: even when two seller–buyer pairs are equidistant, their realized sales can differ substantially depending on the intermediary's location.

While informative, the bivariate plots abstract from multilateral resistance, market size, and institutional linkages, and they necessarily exclude triads that never engage in trade. As a result, bilateral distances alone understate the true spatial frictions in intermediated trade. To address this limitation, Section 3 estimates a triadic gravity specification with a rich set of fixed effects, allowing us to isolate the marginal role of each distance component.

Average markup per buyer or seller. Figure 4 plots the average intermediary markup against the number of counterparties on each side of the market. Panel (a) relates the average markup per buyer to the number of buyers served by intermediaries, while panel (b) shows the corresponding relationship for sellers. Markups are computed as $\frac{\text{sales}}{\text{sales-profit}}$, which captures the price-cost ratio.

A clear inverse relationship emerges: countries with more buyers or sellers exhibit lower average markups per relationship. This pattern suggests that intermediaries operating in thicker markets, where competition is stronger and transactions are more frequent, charge smaller margins. Conversely, intermediaries linked to only a few trading partners sustain higher markups, reflecting limited competition or higher coordination costs per match.

These findings are consistent with competition in thick markets. When more buyers and sellers

interact, intermediaries face stronger competition for matching opportunities, compressing their margins. The negative relationship between network size and markups echoes findings in Ahn et al. (2011) and Akerman (2018): intermediaries earn thinner rents precisely where markets are most active, as the abundance of potential counterparties erodes informational rents.

Market size on both sides of the ledger. Log country-level sales at origins and destinations are strongly positively correlated: countries that are large importers in this network also tend to be large exporters, echoing evidence that market access costs operate symmetrically for buyers and sellers (Ahn et al., 2011; Abel-Koch, 2013). The largest markets—Korea, Singapore, the United States, China, Vietnam, and India—simultaneously serve as major buyers, major suppliers, and hosts of intermediary offices.

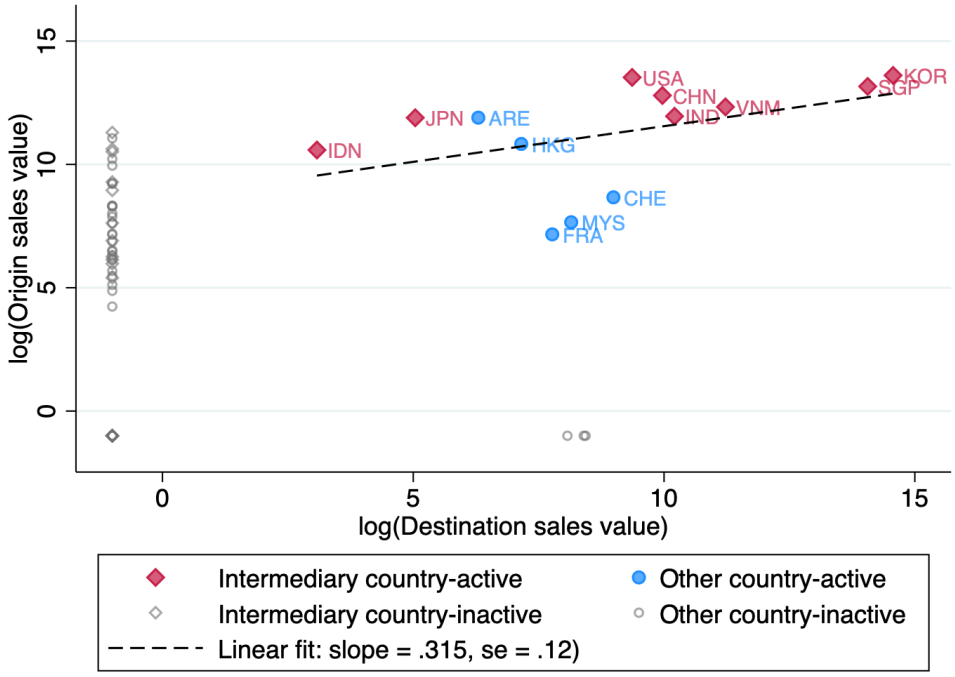


Figure 5: Logged country-level sales at origins and destinations. Red diamonds: countries hosting an intermediary office that also appear on both sides. Blue circles: active markets without an office. Gray markers: countries active on only one side.

3 Intermediary Gravity

Before presenting the estimates, two results previewed by the theory in Section 4 are worth flagging. First, once all three distance legs are controlled for, the conventional negative bilateral elasticity reappears—a reversal that Section 4 rationalizes through the mode-switching mechanism. Second, the positive coefficient on seller-hub distance (d_{om}) is an additional empirical regularity; as noted above, the structural model is silent on this margin, and we interpret it as reflecting supply-side scale economies at the hub.

Our analysis begins with a dyadic gravity framework that captures the role of geography between buyers and sellers, conditional on the intermediary's connection in the seller's country. This specification incorporates the canonical effect of physical distance between origin and destination, while controlling for a rich set of fixed effects. We then extend this framework to a triadic gravity equation that explicitly accounts for the geography of intermediaries and the additional trade frictions they introduce.

3.1 Dyadic and triadic gravity

Specifications. Let $Y_{ijm,g,t}$ denote the outcome of interest for transactions between seller i (in country o) and buyer j (in country d), intermediated by intermediary m, in month t for product g. We estimate the following gravity specification:

$$Y_{ijm,g,t} = \beta_0 + \boldsymbol{\beta}'\mathbf{d}_{od} + FE_{o,year} + FE_{d,year} + FE_g + \epsilon_{ijm,g,t}. \tag{1}$$

Here, $\boldsymbol{\beta}'$ captures the effects of bilateral trade frictions, including the geographic distance as well as common language, colonial ties, and shared borders between the seller and the buyer.

Because most seller-buyer pairs are inactive (as suggested by Figure 1), we rely on origin-year and destination-year fixed effects, $FE_{o,year}$ and $FE_{d,year}$, rather than firm-level or monthly fixed effects. This choice reflects both data coverage and the level at which key shocks operate. Using country-year fixed effects allows us to absorb time-varying factors that affect all exporters or all importers within a country, such as exchange rates, trade policy changes, or demand conditions. Monthly fixed effects would be too granular given that many firm-pair observations are zero or

missing in individual months, while country-year effects balance flexibility with statistical precision. Product fixed effects, FE_g, further control for market conditions specific to each product, as highlighted in Table A1 in the appendix.

Following the convention in the literature on intermediaries as wholesalers, we restrict the dyadic sample to cases where the intermediary office is located in the seller's country, that is $m = o$. Importantly, all transactions in our data are intermediated by the trading house; the direct/indirect distinction refers to whether shipments are routed through a third-country hub, not to the presence or absence of an intermediary.

While the specification in eq. (1) is widely used to model international trade flows, it abstracts from the presence of foreign intermediaries whose locations introduce additional geographical frictions between exporters and importers.

A key innovation of our framework is that it explicitly incorporates the geography of intermediaries, enabling us to account for the distance between the intermediary and both the exporter and the importer. To capture this triadic structure of trade relationships, we extend the conventional bilateral gravity equation as follows:

$$Y_{ijm,g,t} = \beta_0 + \boldsymbol{\beta}_1' \mathbf{d}_{od} + \boldsymbol{\beta}_2' \mathbf{d}_{om} + \boldsymbol{\beta}_3' \mathbf{d}_{md} + FE_{o,year} + FE_{d,year} + FE_{m,year} + FE_g + \epsilon_{ijm,g,t}. \quad (2)$$

Each vector $\mathbf{d}_{uv}$ $(u, v \in \{o, d, m\})$ includes the same set of bilateral characteristics defined in the dyadic specification. The coefficients $(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3)'$ therefore capture how these triadic geographic factors jointly shape trade outcomes.

Identification of these coefficients relies on the assumption that unobserved shocks to the outcome variable, $\epsilon_{ijm,g,t}$, are orthogonal to these bilateral characteristics, conditional on the included fixed effects. This assumption is standard in gravity models, as geographical and historical linkages are largely predetermined and exogenous to short-run trade shocks.

Theoretical motivation. The three-distance specification in eq. (2) is not ad hoc: Proposition 2 in Section 4 shows that, under Pareto productivity and CES demand, bilateral exports factorize into iceberg costs on each leg multiplied by a band-specific market-access term. The log-linearization of this expression at the triad level yields precisely the three-distance estimating equation above,

with leg-specific coefficients mapping to $(1-\sigma)$ times the relevant iceberg components. Formal derivation is deferred to Section 4; the regressions below are best read as reduced-form triadic correlations whose structural interpretation is established there.

Table 3: PPML estimates of dyadic and triadic gravity models

	Quantity		Sales Value		Markup	
	(1)	(2)	(3)	(4)	(5)	(6)
$\log(\text{distance}_{od})$	-0.010	-0.221**	0.131***	-0.050	0.150***	-0.046
	(0.103)	(0.090)	(0.042)	(0.051)	(0.011)	(0.041)
$\log(\text{distance}_{om})$		0.309***		0.077		0.126*
		(0.057)		(0.051)		(0.066)
$\log(\text{distance}_{md})$		-0.045		-0.040		0.297***
		(0.074)		(0.055)		(0.066)
Origin-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Intermediary-year FE	No	Yes	No	Yes	No	Yes
Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-Squared	0.409	0.482	0.368	0.493	0.370	0.464
Log-likelihood	-3.06e+06	-1.54e+07	-8.02e+05	-4.80e+06	-3489.166	-3.30e+04
Observations	1569	12771	1493	12559	1570	12839

Notes: This table reports Poisson pseudo-maximum likelihood (PPML) estimates of gravity models. Column headers indicate the dependent variables. Columns (1), (3), and (5) present results from the dyadic gravity specification, while columns (2), (4), and (6) show the corresponding triadic regression results. Each regression includes bilateral controls for common language, colonial ties, and shared borders, though their coefficients are omitted for brevity. Robust standard errors (in parentheses) are clustered at the origin-destination level for dyadic models and at the origin-intermediary-destination level for triadic models. Statistical significance is denoted by ***, **, and * at the 1%, 5%, and 10% levels, respectively.

Results. Table 3 reports the Poisson pseudo-maximum-likelihood (PPML) estimation results from the dyadic and triadic specifications. We employ PPML because it accommodates zero trade flows and remains consistent in the presence of heteroskedasticity, both of which are common features of disaggregated trade data (Silva and Tenreyro, 2006). Columns (1), (3), and (5) present results from the dyadic wholesaler gravity estimation, while columns (2), (4), and (6) report the corresponding triadic regression results that incorporate the geography of intermediaries.[7] The dependent variables are, respectively, the traded quantity, the sales value, and the intermediary

[7]The sample expands substantially between the dyadic and triadic columns—from 1,569 to 12,771 observations—because the triadic specification includes all intermediary locations, not only cases where $m = o$; see Section A.6 for details.

markup, defined as the price–cost ratio.[8] Each regression also includes bilateral controls for common language, colonial ties, and shared borders, but the corresponding estimates are suppressed to conserve space. Robust standard errors are clustered at the origin–destination level for the dyadic models and at the origin–intermediary–destination level for the triadic models.

For traded quantity, the dyadic specification shows that the distance between the exporter and importer is statistically insignificant, suggesting a potential selection effect. Because the dyadic framework effectively treats intermediaries as domestic wholesalers ($m = o$), the sample already reflects sellers who have optimized their choice of local partners. This pre-selection likely dampens the measured impact of bilateral distance.

Once the intermediary's geography is explicitly incorporated into the triadic specification, the conventional distance elasticity re-emerges: a 1% increase in distance is associated with roughly a 0.2% decline in traded quantities. This result confirms that greater geographic separation between origin and destination reduces trade once the intermediary's position is properly accounted for. At the same time, the distance between the exporter and the intermediary (distance_{om}) is positive and significant, indicating that intermediaries farther from their suppliers handle larger shipment volumes. This pattern suggests that intermediaries act as aggregation hubs, pooling products from more remote sellers to exploit sourcing scope and economies of scale. In contrast, the intermediary-buyer distance (distance_{md}) has no significant effect on traded quantities, implying that shipment size is primarily determined by supply-side consolidation rather than proximity to buyers.

While exporter-importer distance is insignificant for traded quantities, it becomes positive and significant for sales values–a pattern that again points to selection. When distance acts as a barrier to direct shipment, high-value or specialized transactions are likely to survive, producing positive sales elasticities despite limited overall volume responses. In our estimates, the dyadic model in column (3) shows a positive and significant elasticity of exporter–importer distance, while in the triadic version (column 4) this effect disappears once intermediary distances are introduced. Neither the exporter–intermediary nor the intermediary–buyer distance significantly affects sales values, suggesting that geography primarily affects traded quantities rather than prices.

The markup regressions in columns (5) and (6) reveal distinct patterns. In the dyadic model, in-

[8]Throughout this paper, "markup" refers to this trading-house intermediation margin, not the CES variable markup $\sigma/(\sigma-1)$ that appears in the theoretical model.

termediaries charge significantly higher markups for longer-distance trades, consistent with greater informational or coordination frictions. However, once the intermediary's geography is included, the direct origin-destination effect vanishes, and the markup increases instead with the intermediary's distance to the destination. In contrast, the exporter-intermediary distance has a smaller positive effect. These results suggest that intermediaries located farther from buyers command higher margins, potentially reflecting both greater market power and higher service costs associated with reaching distant buyers. In particular, these margins appear to incorporate last-mile fixed costs such as customs clearance, certification, and after-sales services. These triadic distance patterns are robust to controlling for mode-selection: Section A.7 adds the structural Switch statistic (derived in Section 4.7) to the triadic specification, and the leg-specific distance coefficients are qualitatively unchanged, confirming that the triadic gradients capture intensive-margin variation rather than selection composition.[9]

These shifts replicate and extend the benchmark findings in Bernard et al. (2010). Similar to their evidence on U.S. wholesalers, our dyadic results show a flattened or even reversed bilateral elasticity, suggesting that intermediaries sustain trade flowing where direct shipments would otherwise be prohibitively costly. The triadic decomposition clarifies the mechanism: once the transaction's routing is made explicit, the conventional negative bilateral elasticity reappears, and the distance premia can be attributed to supply-side economies of scale and destination-side fixed costs.

Our results also align closely with other studies on intermediary trade. The positive markup elasticity on the intermediary-buyer leg echoes Ahn et al. (2011), who show that Chinese brokers earn rents when destination-side entry barriers are high. The muted quantity response to bilateral distance and the positive seller-intermediary coefficient parallel the French evidence in Bernard et al. (2015). The markup premium on longer last-mile legs is consistent with the destination-side friction mechanism in Ahn et al. (2011) and Bernard et al. (2015), who document that intermediaries earn higher rents when buyers face greater entry barriers or geographic frictions; Felbermayr and Jung (2011) complement this by showing that weak contract enforcement raises intermediary use, though

[9]Table 3 uses a restricted dyadic sample in which intermediaries are located in the seller's country, which may introduce selection. Section A.6 reports a full-sample dyadic robustness check that aggregates flows across all intermediary locations. The main distance patterns are stable, though the sales-value coefficient becomes negative once the sample restriction is removed.

they do not predict markups rise with buyer distance. Consistent with the model's characterization of last-mile fixed costs in Section 4, the markup gradient with respect to intermediary-buyer distance persists after controlling for origin–destination distance and product composition. Finally, the strong sourcing-distance effect on shipment size complements Ganapati (2025), where "super-star" wholesalers exploit scope economies by aggregating remote suppliers; his setting is domestic U.S. distribution rather than international transshipment, but the supply-side pooling mechanism is qualitatively analogous. We note that our theoretical model in Section 4 treats the intermediary wedge γ_m as an exogenous primitive and is therefore silent on the *location* of intermediaries relative to sellers; the positive d_{om} coefficient is an empirical regularity whose full structural account—potentially formalizing the scope-economies mechanism of Ganapati (2025) in an international setting—we leave for future work.

Taken together, these results indicate that standard dyadic wholesaler gravity models reveal how brokered flows differ from direct trade, whereas the triadic framework explains why. By disentangling the three legs of trade, the triadic specification restores a conventional bilateral elasticity, attributes volume gains to supply-side scale economies, and links markup premia to destination-side fixed costs. This distinction refines both policy interpretation—whether interventions should target logistics efficiency at the origin or regulatory streamlining at the destination—and the future modeling of intermediated trade.

Scope of inference: a structural case study. Because all transactions flow through a single trading house, readers should interpret this paper as a *structural case study*: we use one firm's complete transaction network as a controlled laboratory for measuring triadic frictions with precision unavailable in aggregate customs data. Two aspects of inference follow directly. First, the *mechanism*—that the presence of an indirect routing option attenuates bilateral trade elasticities and that intermediary geography matters independently—is expected to generalize: the qualitative gravity patterns replicate regularities documented by Blum et al. (2010) using Chilean customs data and by Ahn et al. (2011) using Chinese broker data. Second, *quantitative magnitudes* are specific to this firm's network and should be read as estimates for this intermediary rather than as industry-wide moments. We report all results with this scope clearly in mind; claims about "standard gravity overstating responses" are calibrated on this network's parameters and stated as

illustrative of the mechanism, not as universal benchmarks.

A second dimension of external validity concerns whether the 25 office locations reflect endogenous placement in already high-trade markets, which would inflate hub-location fixed effects and potentially bias the d_{om} coefficient. We test this in Section A.9: a balance test shows that office countries do not have significantly higher 2018 trade intensity than non-office countries ($p = 0.20$ for seller-side, $p = 0.91$ for buyer-side), a logit regression finds that 2018 trade volumes do not predict subsequent office presence, and a 2018-only subsample regression shows that the triadic distance patterns are present before any within-sample office changes could have occurred (we note, however, that in the 2018 cross-section the seller-hub coefficient d_{om} turns negative and insignificant, which we attribute to the dramatically reduced sample size in the first year of data rather than a reversal of the pooled pattern, though we cannot rule out a time-trend component in the positive pooled estimate). Taken together, this evidence is consistent with office placement reflecting regulatory, logistical, or relationship factors rather than selection into already-active markets.

3.2 Gravity for firm-to-firm transactions

The previous analysis examines how intermediated trade varies in quantity, value, and profitability across geography. We now turn to the formation of firm-to-firm connections, focusing on how geographic distance affects the number of transactions intermediaries facilitate. This analysis highlights how intermediaries help exporters penetrate foreign markets by establishing and maintaining trading links (Ahn et al., 2011).

We estimate a gravity model at the country level, where the dependent variable is the number of firm-to-firm transactions ($N_{odm,g,t}$) for product g from origin o to destination d through hub m in month t.[10] As before, the dyadic specification (column 1) restricts the sample to cases in which the intermediary and the seller are located in the same country, while the triadic specification (column 2) introduces two additional distance terms: between the origin and the intermediary (distance_{om}) and between the intermediary and the destination (distance_{md}).

Table 4 reports the PPML estimates. In the dyadic model, the number of firm-to-firm transactions between origin and destination shows no systematic relationship with bilateral distance. Once

[10] We analyze the country level, as this outcome reflects aggregate cross-market connectivity rather than firm-level behavior.

Table 4: PPML estimates for the number of firm-to-firm transactions

Dep. Var.: $N_{odm,g,t}$	(1)	(2)
log(distance$_{od}$)	-0.003	0.026
	(0.016)	(0.030)
log(distance$_{om}$)		-0.070**
		(0.033)
log(distance$_{md}$)		-0.062*
		(0.033)
Origin-year FE	Yes	Yes
Destination-year FE	Yes	Yes
Intermediary-year FE	No	Yes
Product FE	Yes	Yes
Pseudo R-Squared	0.450	0.385
Observations	528	5160

Notes: The dependent variable is the number of firm-to-firm transactions. Bilateral controls are included but not reported. Robust standard errors (in parentheses) are clustered at the origin–destination level in column (1) and at the origin–intermediary–destination level in column (2). Statistical significance is denoted by ***, **, and * at the 1%, 5%, and 10% levels, respectively.

the intermediary's geography is incorporated, however, a clear pattern emerges: both the origin-intermediary and intermediary-buyer distances enter negatively and significantly. A ten-percent increase in either leg reduces the number of matches by roughly 0.6–0.7 percent. The estimate of the direct origin-destination distance remains small and insignificant, indicating that link formation depends more on the intermediary's position than on the total trade route length.

Overall, the results suggest that matching becomes rarer when intermediaries are geographically distant from either the exporter or the buyer, especially from the seller. Combined with the earlier findings in Table 3, where greater seller-intermediary distance is associated with larger shipment volumes, the evidence points to a consistent pattern: distant plants establish fewer intermediary relationships, but those that do tend to trade in larger quantities.

4 A Model of Multi-country Intermediaries

We show that frictions along intermediary routes play a crucial role in shaping international trade flows. To account for the triadic gravity patterns documented above, we develop a parsimonious framework in which exporters choose between direct shipment and hub-mediated routes. The model nests the standard dyadic case, generates a triadic market-access term that maps directly into our estimating equations, and delivers a one-dimensional sufficient statistic that nets out selection at the boundary between the direct and the cheapest indirect route. This allows us to isolate the intensive-margin response of bilateral frictions and interpret the resulting trade elasticities in environments with intermediation.

4.1 Primitives

Countries are indexed by origin o, destination d, and a set of potential intermediary hubs $m \in \{1, \ldots, M\}$. Consumers in d have CES utility

$$U_d = \left(\int_{\omega \in \Omega_d} q_d(\omega)^{(\sigma-1)/\sigma} \, d\omega \right)^{\sigma/(\sigma-1)}, \qquad \sigma > 1.$$

With total expenditure Y_d, demand for variety ω and the ideal price index P_d are

$$q_d(\omega) = \left(\frac{p_d(\omega)}{P_d}\right)^{-\sigma} \frac{Y_d}{P_d}, \qquad P_d = \left(\int_{\omega \in \Omega_d} p_d(\omega)^{1-\sigma} \, d\omega \right)^{1/(1-\sigma)}.$$

Production uses only labor, and each firm supplies a differentiated variety. We treat destination expenditure Y_d as exogenous throughout (partial equilibrium); wages are symmetric and normalized to one, $w_o = w_d = 1$. Each firm draws productivity $\varphi \in [1, \infty)$ from a Pareto distribution $G(\varphi) = 1 - \varphi^{-\theta}$ with $\theta > \sigma - 1$ (Eaton et al., 2011; Melitz, 2003; Chaney, 2008). Domestic fixed operating cost is normalized to zero, and entry requires a sunk cost $f^e > 0$. Marginal cost at the factory gate is $c_o(\varphi) = 1/\varphi$, which implies an f.o.b. price $p_o(\varphi) = \mu/\varphi$, where $\mu \equiv \sigma/(\sigma - 1)$.

Mapping hubs to sorted modes. For a given origin-destination pair (o, d), an exporter chooses a trade mode $k \in \{0, 1, \ldots, M\}$. The index $k = 0$ denotes the direct route.[11] The indices $k \in \{1, \ldots, M\}$ correspond to the M physical intermediary hubs, re-labeled such that they are ordered by increasing composite iceberg cost. Thus, $k = 1$ refers to the route via the most cost-efficient hub (the "cheapest indirect mode"), $k = 2$ to the second-cheapest, and so forth. In the derivation that follows, we refer exclusively to mode k. This re-indexing allows us to characterize the selection mechanism using a single cut-off on a monotone ladder. Let $m(k)$ denote the specific intermediary country corresponding to the sorted mode index k.

4.2 Trade modes and costs

Mode k is a routing-service bundle characterized by a composite iceberg cost $\tau_{od,k}$ and a destination-specific fixed cost $f_{od,k}$. The composite trade costs are defined as

$$(\tau_{od,k}, f_{od,k}) = \begin{cases} (\tau_{od},\ f_{od}) & \text{if } k = 0 \quad \text{(Direct)}, \\ (\gamma_m \tau_{om} \tau_{md},\ f_{md}) & \text{if } k \geq 1 \text{ and } m = m(k). \end{cases}$$

[11] "Direct" refers to the *routing structure*—bilateral cost τ_{od} without third-country transit—not to the absence of intermediation. Every transaction in our data is arranged by the trading house. This differs from the direct/intermediated dichotomy in Bernard et al. (2010), where "direct" means the manufacturer exports without a wholesaler. In our framework, $m = o$ (intermediary in the seller's country) corresponds to the direct mode, while $m = d$ with $m \neq o$ remains an indirect (hub-routed) transaction.

Here, τ_{om} and τ_{md} denote the physical transport costs from the origin to the specific hub $m = m(k)$ and from that hub to the destination, respectively. The intermediary wedge $\gamma_m \geq 1$ and lane fixed cost f_{md} are exogenous lane-level primitives. We interpret (γ_m, f_{md}) as a posted two-part pricing schedule in a competitive intermediation sector. Each hub m applies an iceberg overlay γ_m to cover services such as financial intermediation, regulatory compliance, and risk management.[12]

For a firm in o exporting via mode $k \geq 1$ (mapped to hub m), the buyer in d pays a delivered price $p_{od,k}(\varphi) = \mu\, \gamma_m \tau_{om} \tau_{md} / \varphi$. Competitive pricing in the intermediation sector implies that expected variable receipts and fixed charges cover the intermediary's total costs. For any mode $k \in \{0, 1, \ldots, M\}$, the delivered marginal cost and consumer price in d are

$$c_{od,k}(\varphi) = \tau_{od,k}\, c_o(\varphi) = \frac{\tau_{od,k}}{\varphi}, \qquad p_{od,k}(\varphi) = \mu\, c_{od,k}(\varphi) = \mu\, \frac{\tau_{od,k}}{\varphi}.$$

Under CES demand, revenue and variable profits for a firm serving d through mode k are

$$r_{od,k}(\varphi) = A_d\, \tau_{od,k}^{1-\sigma}\, \varphi^{\,\sigma-1}, \qquad \pi_{od,k}(\varphi) = \frac{1}{\sigma}\, r_{od,k}(\varphi),$$

where

$$A_d \;\equiv\; \frac{Y_d}{P_d^{\sigma-1}} \Big(\frac{\sigma}{\sigma - 1} \Big)^{1-\sigma}.$$

Subtracting the mode-specific fixed cost $f_{od,k}$ gives operating profits:

$$\Pi_{od,k}(\varphi) = \frac{1}{\sigma}\, A_d\, \tau_{od,k}^{1-\sigma}\, \varphi^{\,\sigma-1} - f_{od,k}. \tag{3}$$

We impose two restrictions on variable and fixed trade frictions: an *inverse-ordering* condition that delivers pairwise single crossing, and a *curvature* condition that yields a strict global ordering of all adjacent switch points (a monotone ladder) and simplifies comparative statics.

Assumption 1 (Inverse ordering)**.** *For each* (o, d)*:*

[12] The multiplicative structure $\gamma_m \tau_{om} \tau_{md}$ captures cost compounding along the intermediary chain. Fally and Hillberry (2018) develop a Coasian model of international production chains in which sequential-stage iceberg costs compound in precisely this manner; our intermediated trade setting maps naturally onto their framework.

(a) Sorting by composite variable cost. Modes are indexed so that

$$\tau_{od,0} \;<\; \tau_{od,1} \;<\; \cdots \;<\; \tau_{od,M}.$$

(b) Normalized fixed-cost monotonicity (NFCM). Normalized fixed fees are weakly decreasing:

$$\frac{f_{od,0}}{\tau_{od,0}^{1-\sigma}} \;\geq\; \frac{f_{od,1}}{\tau_{od,1}^{1-\sigma}} \;\geq\; \cdots \;\geq\; \frac{f_{od,M}}{\tau_{od,M}^{1-\sigma}}, \qquad \textit{or equivalently,} \qquad f_{od,0}\,\tau_{od,0}^{\sigma-1} \;\geq\; \cdots \;\geq\; f_{od,M}\,\tau_{od,M}^{\sigma-1}.$$

Lower per-unit cost modes (smaller $\tau_{od,k}$) therefore carry higher fixed fees (larger $f_{od,k}$).[13] Equivalently, profit-line slopes $(A_d/\sigma)\tau_{od,k}^{1-\sigma}$ decrease in k, while intercepts $-f_{od,k}$ increase. This matches a two-part pricing schedule for intermediaries: lower marginal costs (low $\tau_{od,k}$) are accompanied by higher fees for membership, compliance, and certification (high $f_{od,k}$). Direct exporting saves the most on variable costs but incurs the highest fixed costs.[14]

Notice that $\tau_{od,0} \;\leq\; \tau_{od,1}$, so mode $k = 0$ (direct exporting) is (weakly) the least-cost "top band," providing the benchmark for comparison. Because modes are indexed so that $k = 1$ is the cheapest indirect alternative, the $(0 \leftrightarrow 1)$ indifference boundary moves monotonically with the spoke cost τ_{od} and nearby mode comparisons can be expressed relative to mode 0 as the baseline.

Assumption 1 guarantees *pairwise* single crossing of all mode-specific profit lines: for any two modes k and k', the corresponding profit functions intersect exactly once. However, this condition alone does not impose a *global* ordering of all adjacent switch points. Because the gaps $\{f_{od,k-1} - f_{od,k}\}$ and $\{\tau_{od,k-1}^{1-\sigma} - \tau_{od,k}^{1-\sigma}\}$ are unrestricted, the adjacent cutoffs may interlace. The next assumption imposes a mild shape restriction that rules out such interlacing and yields a strict ladder of cutoffs.

Assumption 2 (Curvature of the fee frontier)**.** *Define the adjacent secant slopes in the* $(\tau^{\,1-\sigma}, f)$

[13] Because observed outcomes depend only on the upper envelope of mode-specific profit lines, one can prune dominated modes (weakly higher τ and higher f) and relabel the survivors so that lower per-unit cost (higher slope in $\varphi^{\sigma-1}$ space) is paired with higher fixed fee. Thus, the sorting clause is, without loss of generality, applied origin–destination by origin–destination.

[14] This variable-cost–fixed-cost tradeoff mirrors the proximity-concentration tradeoff of Helpman et al. (2004): high-productivity firms choose the low-τ / high-f option (direct exporting, analogous to FDI), while lower-productivity firms choose the high-τ / low-f option (indirect routing, analogous to arm's-length exports). We generalize the two-option HMY structure to an ordered ladder of $M + 1$ modes.

plane by

$$R_{od,k} \equiv \frac{f_{od,k-1} - f_{od,k}}{\tau_{od,k-1}^{1-\sigma} - \tau_{od,k}^{1-\sigma}}, \qquad k = 1, \ldots, M.$$

Then the secant slopes satisfy $R_{od,1} > R_{od,2} > \cdots > R_{od,M}$.

After pruning dominated modes (those with weakly higher $\tau_{od,k}$ and higher $f_{od,k}$), the ordering specified in Assumption 1(a) is purely a labeling choice. A convenient sufficient condition for Assumption 2 is that the sequence $\{\tau_{od,k}^{1-\sigma}\}$ forms a geometric progression with common ratio in $(0, 1)$, and the fixed-cost differences $\{f_{od,k-1} - f_{od,k}\}$ are weakly decreasing. Intuitively, Assumption 1 imposes a negative cross-sectional relationship between composite marginal costs and fixed fees, while Assumption 2 strengthens this by requiring *diminishing* fixed-cost savings per unit of marginal-cost improvement as one moves to successively higher-cost (larger k) modes. This curvature (a convex fee frontier in $\tau^{1-\sigma}$) is not required for the existence of equilibrium, but it greatly simplifies comparative statics and the elasticity and welfare results that follow.

4.3 Choice of mode

From eq. (3), the zero-profit cut-off for mode k (treating the mode in isolation) is

$$\varphi_{od,k}^{*} = \left[\frac{\sigma\, f_{od,k}}{A_d\, \tau_{od,k}^{1-\sigma}}\right]^{1/(\sigma-1)}.$$

This cutoff is increasing in both the fixed cost $f_{od,k}$ and the composite iceberg cost $\tau_{od,k}$ (since $1 - \sigma < 0$), and decreasing in destination market size A_d. Thus, more expensive modes, whether in terms of fixed or variable costs, require higher productivity to break even.

Although zero-profit cutoffs determine whether a mode can ever be used, the allocation across modes in equilibrium is governed by *adjacent indifference* cutoffs. Under Assumption 1, modes are ordered so that each adjacent pair $(k, k+1)$ exhibits a unique indifference productivity:

$$\varphi_{od}^{(k \leftrightarrow k+1)} = \left[\frac{\sigma\,(f_{od,k} - f_{od,k+1})}{A_d\big(\tau_{od,k}^{1-\sigma} - \tau_{od,k+1}^{1-\sigma}\big)}\right]^{\frac{1}{\sigma-1}}. \tag{4}$$

Lower-productivity firms prefer the low-fixed/high-τ option (k+1), while higher-productivity firms prefer the high-fixed/low-τ option (k).

It is convenient to define $\tilde{\varphi} \equiv \varphi^{\sigma-1}$, in which case profits are linear in $\tilde{\varphi}$:

$$\Pi_{od,k}(\tilde{\varphi}) = \frac{A_d}{\sigma}\, \tau_{od,k}^{1-\sigma}\, \tilde{\varphi} - f_{od,k}.$$

Under this change of variables, the zero-profit cutoff can be written as

$$\tilde{\varphi}^*_{od,k} = \left(\varphi^*_{od,k}\right)^{\sigma-1} = \frac{\sigma f_{od,k}}{A_d\, \tau_{od,k}^{1-\sigma}},$$

and the adjacent indifference cutoff between modes k and $k+1$ becomes

$$\tilde{\varphi}_{k,k+1} \equiv \frac{\sigma\left(f_{od,k} - f_{od,k+1}\right)}{A_d\left(\tau_{od,k}^{1-\sigma} - \tau_{od,k+1}^{1-\sigma}\right)}.$$

Assumptions 1 and 2 delivers a clean characterization of selection into indirect exporting. In Section B.1, Lemma 2 establishes pairwise single crossing of all profit lines, and Lemma 3 shows that zero-profit cutoffs never bind for interior modes. Proposition 1 then proves that selection is strictly hierarchical: the indifference cutoffs form an ordered ladder, enabling simple comparative statics and clean identification.

Proposition 1 (Monotone cutoff ladder)**.** *Under Assumption 1(a) and Assumption 2, the adjacent indifference cutoffs in $\tilde{\varphi}$ satisfy*

$$\tilde{\varphi}_{0,1} \;>\; \tilde{\varphi}_{1,2} \;>\; \cdots \;>\; \tilde{\varphi}_{M-1,M}.$$

Each mode k occupies the interval

$$\tilde{\varphi} \in \left[\,\tilde{\varphi}_{k,k+1},\ \tilde{\varphi}_{k-1,k}\right), \qquad k = 1, \ldots, M-1,$$

with the direct mode $k = 0$ occupying $[\tilde{\varphi}_{0,1}, \infty)$ and the highest-cost indirect mode $k = M$ occupying $[0, \tilde{\varphi}_{M-1,M})$.

Proof. From (4), $\tilde{\varphi}_{j,j+1} = \frac{\sigma}{A_d} R_{od,j+1}$ for any adjacent pair j. Setting $j = k-1$ gives $\tilde{\varphi}_{k-1,k} = \frac{\sigma}{A_d}\, R_{od,k}$, and setting $j = k$ gives $\tilde{\varphi}_{k,k+1} = \frac{\sigma}{A_d}\, R_{od,k+1}$. Assumption 2 requires $R_{od,k} > R_{od,k+1}$, which implies $\tilde{\varphi}_{k-1,k} > \tilde{\varphi}_{k,k+1}$ for all k. By Lemma 3, zero-profit cutoffs never bind for interior

modes, so equilibrium mode choice is governed entirely by these adjacent indifference cutoffs. The band structure then follows from pairwise single crossing presented in Lemma 2. □

Assumption 1 ensures that mode-specific profit lines are ordered inversely in slope and intercept, which delivers pairwise single crossing. Combined with Lemma 3, this implies that zero-profit cutoffs never bind for interior modes, so equilibrium mode choice is governed entirely by the adjacent indifference cutoffs. Assumption 2 strengthens this local structure into a global ladder: all adjacent cutoffs are strictly ordered ($\tilde{\varphi}_{0,1} > \cdots > \tilde{\varphi}_{M-1,M}$), yielding a clean band structure in which each mode occupies a single, non-overlapping interval of the productivity space. If Assumption 2 fails locally, one can prune or merge adjacent modes whose secant slopes violate monotonicity; the resulting envelope still features pairwise single crossing, and all elasticity bounds continue to hold.

4.4 Equilibrium

Exporters choose, destination by destination, the mode k that maximizes profits, paying the associated fixed costs additively. Under the monotone cutoff ladder in Proposition 1, the productivity space is partitioned into up to M interior intervals plus the terminal direct-export interval for $k = 0$. Free entry pins down the mass of entrants via the zero-expected-profit condition, and the aggregate price index follows the standard Melitz (2003) expression.

Let M_o denote the ex-ante measure of potential entrants in origin o. For each destination d and mode k, the set of productivities that export through k is an interval $[\varphi_{od,k}^{\text{low}}, \varphi_{od,k}^{\text{high}})$, where the endpoints are determined by the ordered indifference cutoffs in Proposition 1.[15]

Since delivered prices satisfy $p_{od,k}(\varphi) = \frac{\sigma}{\sigma-1}\frac{\tau_{od,k}}{\varphi}$, their contribution to the CES price index is

$$
\begin{aligned}
P_d^{1-\sigma} &= \left(\tfrac{\sigma}{\sigma-1}\right)^{1-\sigma} \sum_o \sum_k M_o \, \tau_{od,k}^{1-\sigma} \int_{\varphi_{od,k}^{\text{low}}}^{\varphi_{od,k}^{\text{high}}} \varphi^{\sigma-1} \, g(\varphi) \, d\varphi \\
&= \left(\tfrac{\sigma}{\sigma-1}\right)^{1-\sigma} \sum_o \sum_k M_o \, \tau_{od,k}^{1-\sigma} \, \frac{\theta}{\theta-(\sigma-1)} \Big[(\varphi_{od,k}^{\text{low}})^{\sigma-1-\theta} - (\varphi_{od,k}^{\text{high}})^{\sigma-1-\theta} \Big],
\end{aligned}
$$

and the term involving $\varphi_{od,k}^{\text{high}}$ vanishes whenever $\varphi_{od,k}^{\text{high}} = \infty$ because $\theta > \sigma - 1$.

[15] Specifically, the bounds correspond to the adjacent indifference thresholds defined in eq. (4): $\varphi_{od,k}^{\text{low}} \equiv \varphi_{od}^{(k \leftrightarrow k+1)}$ and $\varphi_{od,k}^{\text{high}} \equiv \varphi_{od}^{(k-1 \leftrightarrow k)}$. For the direct mode ($k = 0$), the upper bound is $\varphi_{od,0}^{\text{high}} = \infty$. For the highest-cost mode ($k = M$), the lower bound is determined by the zero-profit condition $\varphi_{od,M}^{*}$ rather than an indifference condition.

Free entry requires that the expected operating profit equal the entry cost:

$$\sum_{d}\sum_{k}\int_{\varphi_{od,k}^{\text{low}}}^{\varphi_{od,k}^{\text{high}}} \Pi_{od,k}(\varphi)\, g(\varphi)\, d\varphi = f^{e}.$$

Together with labor-market clearing and the price-index expression above, these conditions close the model. Because labor supply is perfectly elastic at wage $w_o = 1$ (workers have access to an outside sector at constant returns), labor-market clearing does not further restrict M_o, which is determined solely by the zero-expected-profit condition above. Two standard benchmarks are nested as special cases: setting $M = 1$ nests a single-hub environment that captures the direct-vs-intermediated trade-off highlighted in Ahn et al. (2011), though through a variable-cost iceberg wedge γ_m rather than the fixed-cost reduction in their framework; letting all indirect fixed costs $f_{od,k} \to \infty$ collapses the model to the canonical Melitz (2003) setting with only direct exporting.

4.5 Triadic gravity, mode shares, and market access

With Pareto productivity $g(\varphi) = \theta\, \varphi^{-\theta-1}$ and $\theta > \sigma - 1$, define the origin-scale term

$$T_o \;\equiv\; M_o \,\frac{\theta}{\theta - (\sigma - 1)}.$$

Under Assumptions 1 and 2, the productivity space in destination d is partitioned into disjoint bands $[\varphi_{od,k}^{\text{low}}, \varphi_{od,k}^{\text{high}})$ for modes $k \in \{0, 1, \ldots, M\}$. For each band, define

$$Z_{od,k} \;\equiv\; \big(\varphi_{od,k}^{\text{low}}\big)^{\sigma-1-\theta} - \big(\varphi_{od,k}^{\text{high}}\big)^{\sigma-1-\theta} \;\geq 0,$$

which depends only on the endogenous cutoffs and hence on the triadic friction vector $\{\tau_{od,k}, f_{od,k}\}$.

Proposition 2 (Triadic gravity and market access)**.** *For any trade mode k with composite iceberg cost $\tau_{od,k}$, bilateral exports from o to d satisfy*

$$X_{od} \;=\; A_d\, T_o \sum_{k=0}^{M} \tau_{od,k}^{1-\sigma}\, Z_{od,k}, \qquad X_{od,k} \;=\; A_d\, T_o\, \tau_{od,k}^{1-\sigma}\, Z_{od,k}.$$

Summing across origins yields the destination market-access index

$$\Phi_d \equiv \sum_o T_o \sum_{k=0}^{M} \tau_{od,k}^{1-\sigma} Z_{od,k},$$

and the CES price index in d can be written as

$$P_d^{1-\sigma} = \left(\tfrac{\sigma}{\sigma-1}\right)^{1-\sigma} \Phi_d.$$

Proof. Revenue of a firm with productivity φ exporting through mode k is $r_{od,k}(\varphi) = A_d \tau_{od,k}^{1-\sigma} \varphi^{\sigma-1}$. Integrating over $[\varphi_{od,k}^{\text{low}}, \varphi_{od,k}^{\text{high}})$ and multiplying by M_o yields $X_{od}^k = A_d \tau_{od,k}^{1-\sigma} M_o \frac{\theta}{\theta-(\sigma-1)} \big[(\varphi_{od,k}^{\text{low}})^{\sigma-1-\theta} - (\varphi_{od,k}^{\text{high}})^{\sigma-1-\theta}\big]$, which equals $A_d T_o \tau_{od,k}^{1-\sigma} Z_{od,k}$. Summing over k gives X_{od}, and summing across origins produces the market-access identity by direct comparison with the CES price-index expression. □

Mode shares are purely extensive-margin objects,

$$S_{od,k} \equiv \frac{X_{od,k}}{X_{od}} = \frac{\tau_{od,k}^{1-\sigma} Z_{od,k}}{\sum_{k'=0}^{M} \tau_{od,k'}^{1-\sigma} Z_{od,k'}}, \qquad \sum_{k=0}^{M} S_{od,k} = 1.$$

They depend only on triadic frictions and the fixed-cost schedule (via cutoffs)–not on expenditure Y_d, the price index P_d, or the origin-scale parameter T_o. Thus, mode choice is solely a function of the route's geometry and the fee structure.

For any perturbation to the lane-specific parameters $\{\tau_{od,k}\}$, bilateral exports satisfy the decomposition

$$d \ln X_{od} = \underbrace{\sum_{k=0}^{M} S_{od,k} \, (1-\sigma) \, d \ln \tau_{od,k}}_{\text{intensive}} + \underbrace{\sum_{k=0}^{M} S_{od,k} \, d \ln Z_{od,k}}_{\text{extensive (selection)}}.$$

The first term is the direct elasticity of each mode's variable trade cost weighted by its share. The second captures how productivity cutoffs shift in response to the same shock. Under our monotone cutoff ladder, only the boundary between direct exporting and the cheapest indirect mode moves with τ_{od}; all non-adjacent bands remain locally invariant. This property is what later permits a closed-form elasticity expression (see Proposition 3).

4.6 Trade elasticities

The presence of multiple intermediary routes alters the link between bilateral trade frictions and aggregate exports in two distinct ways. The key distinction is whether indirect modes introduce additional physical legs beyond the direct route. When all modes share the same iceberg factor, multiple modes affect only the level of exports. When indirect routes introduce their own physical leg, selection across modes dampens the canonical Melitz elasticity.

Case 1 (path–invariant iceberg). Multiple modes shift the level of bilateral exports through the extensive margin but leave the log-elasticity untouched.[16]

Suppose the composite iceberg factor is identical across all modes (i.e., $\tau_{od,k} = \tau_{od}$). In this case, all indifference cutoffs scale proportionally with τ_{od}, and each band inherits the same elasticity. Extensive–margin reallocation shifts levels only, leaving the bilateral elasticity at the canonical value (Chaney, 2008). Lemma 4 in Section B.2 shows formally that

$$\varepsilon_{od} \equiv \frac{\partial \ln X_{od}}{\partial \ln \tau_{od}} = -\theta.$$

Intuitively, without route-specific physical legs, intermediaries do not change the marginal effect of the direct spoke cost. Selection across modes does not distort the log-response, so one recovers the canonical $-\theta$ elasticity (Chaney, 2008).

Case 2 (physical trans-shipment). Now assume that indirect modes involve their own physical leg (i.e., $\tau_{od,k} = \gamma_m \tau_{om} \tau_{md}$). In this case, the bilateral elasticity becomes strictly attenuated:

$$0 \;\geq\; \varepsilon_{od} \;\geq\; -\theta.$$

The reason is that a change in the direct spoke (o, d) shifts the adjacent boundary between the direct mode $k = 0$ and the cheapest indirect mode $k = 1$. All other modes $k \geq 2$ remain locally unaffected under the monotone cutoff ladder.

Two sufficient statistics characterize the magnitude of this attenuation: (i) Boundary-sensitivity index (BSI) Ψ_{od}: how close the direct mode is to the next-best indirect composite friction, and (ii)

[16] This result does not rely on Assumptions 1 and 2 or any restriction on $f_{od,k}$ beyond $\partial f_{od,k}/\partial \tau_{od} = 0$.

Adjacent-switch weight (ASW) Λ_{od}: how much value lies near the boundary and is marginal with respect to the spoke cost.

The next proposition provides the exact decomposition.

Proposition 3 (Bilateral elasticity with trans-shipment)**.** *Work under Case 2 and Assumptions 1 and 2.*

Define Boundary-sensitivity index (BSI)

$$\Psi_{od} \;\equiv\; \frac{\tau_{od,0}^{1-\sigma}}{\tau_{od,0}^{1-\sigma} - \tau_{od,1}^{1-\sigma}}.$$

With the direct value share $S_{od,k} \equiv X_{od,k}/X_{od}$ and the adjacent indifference cutoffs $\varphi_{0,1}$ and $\varphi_{1,2}$ (in raw productivity space), define Adjacent-switch weight (ASW)

$$\Lambda_{od} \;\equiv\; S_{od,1}\,\frac{\varphi_{0,1}^{-\nu}}{\varphi_{1,2}^{-\nu} - \varphi_{0,1}^{-\nu}},$$

where $\nu \equiv \theta-(\sigma-1) > 0$. (Note: ν here is a local exponent; it is distinct from the gravity coefficient vectors $\boldsymbol{\beta}$ estimated in Section 3.) The bilateral trade elasticity with respect to the direct spoke cost is

$$\varepsilon_{od} \;\equiv\; \frac{\partial \ln X_{od}}{\partial \ln \tau_{od}} \;=\; -(\sigma-1)\,S_{od,0} \;-\; \big[\theta-(\sigma-1)\big]\,\underbrace{\Psi_{od}}_{BSI}\,\big(S_{od,0} - \underbrace{\Lambda_{od}}_{ASW}\big)\,. \tag{5}$$

Under the curvature-ladder structure, only the boundary between $k=0$ and $k=1$ moves with τ_{od}; bands $k \geq 2$ are locally fixed and contribute no additional terms.

Proof. See Section B.3. □

Intuitively, the BSI captures the proximity of the direct route to its nearest indirect alternative—a technological or geometric notion—while the ASW captures how much mass is available to reallocate across that boundary.

The elasticity in Proposition 3 contains two components. The first term, $-(\sigma-1)S_{od,0}$, is the intensive-margin response on the direct mode: the canonical elasticity scaled by the direct export share. The second term is the selection margin. As τ_{od} rises, firms re-route toward the adjacent indirect mode ($S_{od,0}$ falls). This re-routing generates a selection effect that partially offsets the intensive-margin decline, shrinking the aggregate elasticity toward zero.

This decomposition nests the canonical Chaney (2008) elasticity as a limiting case. If the indirect route becomes prohibitively expensive ($\tau_{od,1} \to \infty$), the switching mass vanishes ($\Lambda_{od} \to 0$), the boundary sensitivity approaches unity ($\Psi_{od} \to 1$), and the direct share approaches one. Equation (5) then collapses to $\varepsilon_{od} = -(\sigma - 1) - [\theta - (\sigma - 1)] = -\theta$. Conversely, the presence of a viable intermediary option dampens the extensive-margin loss, strictly attenuating the trade response relative to the single-channel benchmark.

Lemma 1 (Direct share dominates ASW)**.** *Under Assumptions 1 and 2 and Pareto productivity with $\sigma > 1$, the direct export share weakly dominates the adjacent-switch weight: $S_{od,0} \geq \Lambda_{od}$.*

Proof. See Section B.4. □

The inequality $S_{od,0} \geq \Lambda_{od}$ ensures that the selection term in (5) is non-positive: the direct mode captures more trade value than the mass that is marginal to re-routing, so every firm that switches from direct to indirect unambiguously reduces bilateral trade. Moreover, the proof of Lemma 1 yields the exact identity $\Psi_{od}(S_{od,0} - \Lambda_{od}) = S_{od,0}$, which collapses the elasticity to $\varepsilon_{od} = -\theta\, S_{od,0}$. This compact expression drives both the bounds in the following corollary and the counterfactual calculations in Section 5.1.

Corollary 1 (Bounds and limits)**.** *Under the assumptions of Proposition 3, the bilateral elasticity satisfies*

$$0 \;\geq\; \varepsilon_{od} \;\geq\; -\theta.$$

Moreover, $\varepsilon_{od} \to -\theta$ as the indirect option becomes irrelevant ($\tau_{od,0}/\tau_{od,1} \to 0$), and $\varepsilon_{od} \to 0$ as the direct cost approaches the indirect frontier ($\tau_{od,0} \to \tau_{od,1}$).

Proof. See Section B.5. □

The bilateral elasticity depends only on triad-level objects $\{S_{od,0}, \Lambda_{od}, \Psi_{od}\}$, not on destination demand or origin scale. Any improvement to an indirect leg (lowering $\tau_{od,1}$) raises the boundary sensitivity Ψ_{od} and shifts mass away from the direct mode, thereby attenuating $|\varepsilon_{od}|$ by expanding the scope for re-routing.

Remark 1 (Robustness to Assumption 2)**.** *If the curvature condition in Assumption 2 fails and multiple adjacent boundaries co-move with $\tau_{od,0}$, each additional moving boundary adds a non-*

negative selection term to Proposition 3. The attenuation bound in Corollary 1 therefore continues to hold: $\varepsilon_{od} \in [-\hat{\theta} \cdot S_{od,0},\, 0]$. *The formula in Proposition 3 is a lower bound on* $|\varepsilon_{od}|$ *when Assumption 2 fails, so the paper's quantitative estimates of attenuation are conservative in this sense.*

In practice, trade shocks—such as port upgrades or tariffs—often hit a physical link (o, d) rather than a unique commercial pair. Because the leg (o, d) serves as both the direct route for local exporters and the transit route for third-country trans-shipments, a shock to τ_{od} propagates through multiple origins. To capture this, we define the arc elasticity, aggregating responses across all origins traversing the link.

Proposition 4 in Section B.6 formally derives this aggregation. It shows that the arc elasticity is bounded by:

$$\varepsilon_{\cdot d}^{\text{arc}(od)} \in \big[\, -\theta\, S_{\cdot d}^{\text{arc}(od)},\, 0\big],$$

where $S_{\cdot d}^{\text{arc}(od)}$ is the share of destination d's total imports that traverse the arc (o, d). This result confirms that the attenuation mechanism scales up: the aggregate response is strictly smaller than the Melitz benchmark applied to the active arc share.[17]

4.7 Selection into intermediaries and a sufficient-statistics control

Since our data condition on the realized intermediary, estimating intensive-margin elasticities requires controlling for the extensive margin of mode selection. Under the monotone ladder structure defined in Section 4, a change in bilateral costs τ_{od} locally affects the boundary between direct exporting ($k = 0$) and the cheapest indirect mode ($k = 1$). Ignoring this boundary movement biases the estimate of the direct leg coefficient by conflating the intensive-margin response with the reallocation of firms across this boundary.

We derive a scalar sufficient statistic that exactly nets out this selection margin for the direct mode. As detailed in Section C.1, the indifference cutoff $\varphi_{0,1}$ is a power function of the gap in composite variable costs between the direct route and the best available indirect route. We therefore define the selection control term as:

$$\widehat{\mathsf{Switch}}_{od} = \frac{1}{\sigma - 1}\, \ln\Big(\tau_{od,0}^{1-\sigma} - \tau_{od,1}^{1-\sigma}\Big), \qquad \text{where } \tau_{od,1} \equiv \min_{m}(\gamma_m \tau_{om} \tau_{md}).$$

[17] The knife-edge Case 1 with identical iceberg costs across modes ($\tau_{od,k} = \tau_{od}$) shuts down the selection margin entirely, yielding the standard Melitz result $\varepsilon_{\cdot d}^{\text{arc}(od)} = -\theta \cdot S_{\cdot d}^{\text{arc}(od)}$.

Because direct exporting entails lower variable costs ($\tau_{od,0} < \tau_{od,1}$), the term inside the logarithm is strictly positive.[18]

Data construction of the Switch statistic. In the data, $\tau_{od,k}$ is unobserved. We proxy iceberg costs by CEPII great-circle distances, assuming proportionality $\tau_{od,k} \propto d_k^{\rho}$. For the minimum-cost indirect route, $\tau_{od,1} \equiv \min_m(\gamma_m \tau_{om} \tau_{md})$, we (i) set the intermediary-specific wedge $\gamma_m = 1$ for all hubs (hub-level fixed effects in the triadic regression absorb level differences in γ_m), and (ii) take the minimum over all intermediary offices *in the data* for each O-D pair—that is, the hub $m^* \equiv \arg\min_m(d_{om} \cdot d_{md})$ among offices active in the trading house's network. The cost power $\rho > 0$ cancels from the Switch statistic under the log-ratio form: $\widehat{\mathsf{Switch}}_{od} = \frac{\rho(1-\sigma)}{\sigma-1}[\ln d_{od} - \ln(d_{om^*} d_{m^*d})] = -\rho[\ln d_{od} - \ln(d_{om^*} d_{m^*d})]$, so the selection-sensitivity coefficient $\kappa \equiv \frac{\theta-(\sigma-1)}{\sigma-1}$ (formally introduced in (6) below) is absorbed only up to a multiplicative rescaling ($\hat{\kappa} = \kappa_{\text{true}}/\rho$), and the recovered $\hat{\theta} = \hat{\kappa}(\sigma-1) + (\sigma-1)$ is invariant to ρ under the baseline parameterization.

For direct transactions, this term fully absorbs the selection effect. However, identifying the structural parameters presents an econometric challenge. Theoretically, one could estimate both the intensive-margin elasticity $(1-\sigma)$ and the selection sensitivity (κ) simultaneously using the structural iterative procedure we derive in Section C.2. In practice, however, we find that the switch term—which is functionally a log-difference of distances—is highly collinear with the direct distance term used to identify σ. This collinearity leads to numerical instability, preventing the joint identification of both parameters.

The primary source of identifying variation in $\hat{\kappa}$ is cross-sectional hub geometry: the Switch statistic varies across O-D pairs because different pairs face different optimal-hub composite distances ($d_{om^*} \cdot d_{m^*d}$). For most O-D pairs, hub identity is fixed throughout the sample period—the same hub office handles that route in every year—so the variation in Switch is driven by cross-sectional differences in hub geometry rather than by time variation in hub identity. The Russia

[18] The name *Switch* reflects the mode-switching mechanism: the statistic measures the variable-cost advantage of the direct lane over the least-cost hub route. In the distance-proxy implementation, iceberg costs are approximated as $\tau_{od,k} \propto d_k^{\rho}$, so the statistic reduces to a scaled log-ratio of direct versus hub-routing distances. Specifically, with a single hub intermediary, $\widehat{\mathsf{Switch}}_{od} \approx \frac{\rho}{\sigma-1}[\ln d_{od} - \ln(d_{om} \cdot d_{md})]$, i.e., the log of the direct distance minus the sum of the two indirect legs. It is constructed from CEPII great-circle distances matched to exporter–intermediary–importer triads.

episode (Section 5.2), in which a few O-D pairs lost Russia as a feasible hub after 2021, does generate within-pair variation in Switch for those routes; this variation is consistent with the rest but represents a small share of the identifying sample. Regarding the iceberg proportionality assumption: the cost power ρ cancels from the Switch statistic in log-ratio form (as noted above), so misspecification of ρ rescales $\hat{\kappa}$ multiplicatively but does not affect the sign or significance of the attenuation result. The directional finding—that bilateral elasticities are strictly below the Melitz benchmark—is invariant to ρ.

Consequently, we adopt a calibration strategy standard in the trade literature (Head and Mayer, 2014). We fix the intensive-margin elasticity at a baseline value of $\sigma = 4$, consistent with Head and Mayer's meta-analysis of aggregate bilateral gravity estimates (their Table 3), and estimate the selection parameter conditional on it. We note that product-level estimates from Broda and Weinstein (2006) for HS chapters 27–29 (petroleum and petrochemical products) typically fall in the range $\sigma \in [6, 12]$, reflecting the more homogeneous nature of these goods relative to the differentiated-goods average. At those sector-appropriate values, the implied $\hat{\theta}$ ranges from 55 to 89, well above the cross-country benchmarks of $\hat{\theta} \approx 4$–8 from Simonovska and Waugh (2014) and Eaton et al. (2011). We therefore adopt the aggregate gravity meta-analysis range $\sigma \in \{3, 4, 5, 6\}$ as our primary calibration domain; $\sigma = 4$ within this range is conservative in the sense of generating a lower implied $\hat{\theta}$ relative to higher σ values; we report the full sensitivity of $\hat{\kappa}$ and $\hat{\theta}$ to $\sigma \in \{3, 4, 5, 6\}$ in Table C1, and these results extend to higher values in the robustness analysis below. We implement this by treating the distance elasticity as a constrained offset:

$$\ln X_{od,0} = \underbrace{(1-\sigma)\ln \tau_{od,0}}_{\text{Offset}} + \kappa \widehat{\mathsf{Switch}}_{od} + \delta_o + \delta_d + \varepsilon_{od,0}. \tag{6}$$

In this specification, the coefficient $\hat{\kappa}$ captures the sensitivity of the selection margin, theoretically defined as $\frac{\theta-(\sigma-1)}{\sigma-1}$. This allows us to recover the Pareto shape parameter as $\hat{\theta} = \hat{\kappa}(\sigma-1) + (\sigma-1)$. The calibrated offset $-(\sigma-1)$ should not be confused with the reduced-form distance coefficient of approximately -0.2 estimated in Table 3: the latter is a within-fixed-effect partial correlation identified after absorbing origin and destination averages, and is attenuated both by the high-dimensional fixed effects and by indirect routing absorbing part of the bilateral distance response; the former is a structural parameter calibrated from the trade literature and imposed rather than

estimated.

Table 5 presents the estimates from this structural regression restricted to the subsample of direct transactions. Column (1) reports the results for trade quantities. We estimate a selection coefficient of $\hat{\kappa} \approx 2.96$ (s.e. 0.10). This positive, statistically significant coefficient confirms that the option to switch to indirect trade strongly governs the volume of direct exports: when the cost advantage of the direct mode over the indirect alternative is large (i.e., the Switch statistic is high), direct exports are substantially larger; as the cost gap narrows and the Switch falls, firms increasingly shift toward indirect routing.

Using the mapping derived above, this estimate implies a Pareto shape parameter of $\hat{\theta} \approx 11.9$, conditional on the imposed $\sigma = 4$. We emphasize that $\hat{\theta}$ is not independently identified: it is a function of the calibrated σ, and Table C1 in the appendix shows that $\hat{\theta}$ ranges from 5.9 to 29.7 as σ varies from 3 to 6; at sector-appropriate values $\sigma \in [6, 12]$ from Broda and Weinstein (2006), the implied $\hat{\theta}$ is commensurately higher. The appropriate measure of this calibration uncertainty is the cross-row spread in Table C1, not the within-row delta-method confidence intervals, which condition on the calibrated σ and therefore dramatically understate total parameter uncertainty. The directional result—that bilateral elasticities are attenuated relative to $-\hat{\theta}$—is robust throughout: every row in Table C1 yields a positive $\hat{\kappa}$, confirming that the switching mechanism operates regardless of the calibrated σ. For reference, Simonovska and Waugh (2014) estimate $\theta \approx 4$–8 across countries using price data, and Eaton et al. (2011) recover Pareto shape parameters in the range 3–5 for French manufacturing; the selection margin in our intermediated setting naturally implies a higher effective θ because the indirect routing option absorbs part of the cost sensitivity that bilateral models attribute to firm exit. Our preferred estimate at $\sigma = 4$ is consistent with the nature of the petrochemical industry. The products in our sample are highly standardized, low-differentiation commodities. In such markets, firm heterogeneity is relatively low (high θ), meaning that even small differences in trade costs can drive large shifts in mode selection.

Columns (2) and (3) confirm the robustness of this mechanism using sales value and markups as dependent variables. The estimated selection coefficients are quantitatively stable across specifications, ranging from 2.96 to 3.12. Consequently, the recovered structural parameter $\hat{\theta}$ is stable across the quantity and sales-value specifications (11.9 and 12.3, respectively), suggesting that the selection mechanism consistently governs not just physical flows, but also the revenue streams of

the firm. The markup specification does not yield a well-defined $\hat{\theta}$ because markups are not directly interpretable in the Pareto-share aggregation framework.

Table 5: Structural estimates of selection margin (Direct subsample)

	(1) Quantity	(2) Sales Value	(3) Markup
$\widehat{\mathsf{Switch}}_{od}$	2.958***	3.098***	3.116***
	(0.102)	(0.042)	(0.010)
Origin-year FE	Yes	Yes	Yes
Destination-year FE	Yes	Yes	Yes
Intermediary-year FE	No	No	No
Product FE	Yes	Yes	Yes
Pseudo R-Squared	0.409	0.292	0.370
Log-likelihood	-3.06e+06	-987.095	-3489.166
Observations	1569	1493	1570
Clusters (O-D)	20	20	20
Calibrated σ	4	4	4
Recovered θ	11.87	12.30	12.35

Notes: Standard error for $\hat{\kappa}$ in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table reports PPML estimates of the structural selection parameter κ using the specification in Equation (6). The sample is restricted to direct transactions ($iso_o = iso_m$). The coefficient on $\ln \tau_{od,0}$ is constrained to $1 - \sigma = -3$ (implying $\sigma = 4$) using the offset method. Standard errors are clustered at the origin-destination level. With only 20 origin-destination clusters, we apply the $t(G-1) = t(19)$ small-sample correction recommended by Cameron and Miller (2015): the cluster-robust t-statistics for $\hat{\kappa}$ are 28.93, 74.62, and 296.94 for Quantity, Sales Value, and Markup respectively, each exceeding the $t(19)$ 1% critical value of 2.54 by an order of magnitude. We also implement the score-based wild-cluster bootstrap of Cameron et al. (2008) (Rademacher weights, 9,999 replications), but note that this score test has near-zero power in our setting because the doubly-balanced fixed-effect structure forces cluster-level residuals to cancel across origins and destinations; the Wald-based $t(19)$ inference is therefore the appropriate small-sample check.

The key to reconciling the structural parameters with the reduced-form gravity estimates lies in Proposition 3: the aggregate trade elasticity with respect to bilateral costs is not $-\hat{\theta}$ but $-\hat{\theta} \cdot S_{od,0}$, attenuated by the direct export share. Table 6 collects the calibrated, estimated, and implied quantities in one place. Evaluating the model identity $\varepsilon_{od} = -\hat{\theta} \cdot S_{od,0}$ at the reduced-form estimate of -0.221 implies a direct export share of $S_{od,0} \approx 0.019$—consistent with a trading house in which most origin-destination pairs route a small fraction of their total flow directly. The high structural $\hat{\theta} \approx 11.9$ and the low reduced-form elasticity -0.221 are thus two faces of the same fact: selection into indirect routing is strong ($\hat{\theta}$ large) and the share of directly-routed trade is small ($S_{od,0}$ close to zero), so the aggregate elasticity—their product—is small.

The high value of θ reflects the homogeneous nature of petrochemical products. In a market

Table 6: Parameter reconciliation: structural estimates and reduced-form elasticities

Object	Symbol	Value	Source / Interpretation
Panel A: Calibrated			
CES elasticity	σ	4	Head-Mayer meta-analysis; governs intensive margin
Firm-level cost elasticity	$-(\sigma-1)$	-3	Offset in Equation (6); single-firm response to τ, fixed participation
Panel B: Estimated from Switch regression			
Selection coefficient	$\hat{\kappa}$	2.96 (0.10)	Sensitivity of direct exports to cost gap (Table 5)
Pareto shape	$\hat{\theta}$	11.9	$\hat{\kappa}(\sigma-1)+(\sigma-1)$; governs firm selection into direct mode
Panel C: Implied and reduced-form moments			
Model-implied trade elasticity	$-\hat{\theta} \cdot S_{od,0}$	≈ -0.23	Proposition 3; aggregate O–D trade response to τ_{od}
Direct export share (implied)	$S_{od,0}$	≈ 0.019	Backed out from $-\hat{\theta} \cdot S_{od,0} \approx -0.221$
Reduced-form distance elasticity	$\partial \ln X / \partial \ln d_{od}$	-0.221	Table 3 col. (2); within-FE partial correlation

Notes: Standard error for $\hat{\kappa}$ in parentheses. $S_{od,0}$ is not directly estimated but backed out from the Proposition 3 identity at the sample-mean elasticity. The reduced-form and model-implied elasticities are not identical objects—the former conditions on intermediary fixed effects, the latter sums over routing modes—but their quantitative proximity confirms the internal consistency of the structural estimates.

with low differentiation, firms face high elasticity of substitution across transport modes; even minor differences in direct and indirect trade costs trigger significant reallocation of market share.

Crucially, this high sensitivity to costs explains the low observed sensitivity to physical distance. When direct physical frictions increase, the high-θ environment ensures that firms aggressively switch to the indirect network rather than ceasing trade or absorbing the cost. This switching mechanism effectively insulates the aggregate trade volume from bilateral geographic shocks. Consequently, the observed trade elasticity is strictly lower than the canonical Melitz benchmark, as the intermediary network effectively flattens the spatial cost curve.

This finding underscores the importance of accounting for intermediary gravity. In a high-θ environment, omitting the indirect option results in severe misspecification of trade costs. Standard gravity models would interpret a rise in direct costs as a pure destruction of trade, whereas our triadic framework reveals that the market efficiently reroutes these flows through the path of least resistance. Thus, in homogeneous-good sectors, the role of intermediaries in dampening the impact of bilateral geographic frictions is amplified by the high elasticity of substitution.

5 Quantitative Exercises

Our structural framework delivers two quantitative exercises. Exercise (i) quantifies how much standard bilateral gravity overstates the trade-flow response to a bilateral cost shock. Exercise (ii) uses the CES hub-demand formula to bound the implied compliance-cost increase at the Russia hub from the observed volume decline after 2020, illustrating the triadic rerouting mechanism for a documented disruption.

5.1 Attenuation of Bilateral Cost Shocks

Proposition 3 and Corollary 1 imply that the first-order elasticity of total O-D trade with respect to a bilateral spoke cost is $\varepsilon_{od} = -\hat{\theta} \cdot S_{od,0}$, where $S_{od,0}$ is the share of O-D trade value routed through the direct mode. Standard bilateral gravity assumes $\varepsilon = -\hat{\theta}$ for every pair. The gap between the two predictions is determined entirely by $S_{od,0}$, which is directly observable in our data.

The distribution of $S_{od,0}$ across the 128 O-D pairs, reported in full in Table A5 in the appendix, reveals a stark pattern: 84.5% of pairs are *fully indirect* ($S_{od,0} = 0$), meaning they have *zero* first-order sensitivity to a bilateral cost shock under the triadic model. Standard gravity would predict an elasticity of -11.87 for these same pairs. A further 3.2% have low-to-moderate direct shares ($0 < S_{od,0} < 0.75$) with correspondingly attenuated elasticities, and 2.3% have high but incomplete direct shares ($0.75 \leq S_{od,0} < 1$). Only 9.3% of pairs that are fully direct ($S_{od,0} = 1$), where the origin and intermediary country coincide, recover the full standard elasticity. The extreme concentration of direct trade in a small minority of O-D pairs parallels the superstar-exporter pattern documented by Freund and Pierola (2015), where a handful of large exporters account for a disproportionate share of bilateral export value; here, the analogous concentration is in routing mode rather than firm size.

Table 7 documents the pair-level attenuation. Panel A summarises the 109 fully-indirect pairs ($S_{od,0} = 0$): bilateral cost shocks have zero first-order effect on their trade volume. Panel C summarises the 12 fully direct pairs ($S_{od,0} = 1$, where the origin and intermediary country coincide): these recover the full standard elasticity of -11.87. The empirical content of the triadic model lies in Panel B, which lists all seven O-D pairs with intermediate direct shares ($0 < S_{od,0} < 1$). Their implied elasticities range from -0.08 (KOR–SGP, $S_{od,0} = 0.006$, almost entirely intermediated)

to −11.18 (SGP–JPN, $S_{od,0} = 0.942$, near-full direct), tracing out the full attenuation gradient between the two extremes.

Table 7: Standard vs. triadic elasticity by O-D pair

Origin	Destination	$S_{od,0}$	Triadic ε_{od}
Panel A: Fully indirect routes ($S_{od,0} = 0$) — zero bilateral sensitivity			
All 109 O-D pairs with $S_{od,0} = 0$		0.000	0.000
Panel B: Partially-direct routes ($0 < S_{od,0} < 1$) — attenuated elasticity			
KOR	SGP	0.006	−0.08
SGP	USA	0.015	−0.18
SGP	SGP	0.264	−3.13
SGP	KOR	0.293	−3.47
KOR	KOR	0.752	−8.93
SGP	CHE	0.760	−9.03
SGP	JPN	0.942	−11.18
Panel C: Fully direct routes ($S_{od,0} = 1$) — full standard elasticity recovered			
All 12 O-D pairs with $S_{od,0} = 1$		1.000	−11.87

Notes: $\varepsilon_{od} = -\hat{\theta} \cdot S_{od,0}$, with $\hat{\theta} = 11.87$ ($\sigma = 4$). Standard gravity predicts $\varepsilon = -\hat{\theta} = -11.87$ for every pair regardless of routing mode. Panel B lists all O-D pairs with partial direct routing; their elasticities span −0.08 to −11.18, illustrating the attenuation gradient predicted by Proposition 3.

These results deliver the quantitative answer that the structural framework is designed to provide: for the large majority of O-D pairs in this intermediated petrochemical network, a standard gravity model would *overstate* the trade-flow response to a bilateral cost shock by a factor of $1/S_{od,0}$ — ranging from moderate amplification for partially-direct pairs to infinite amplification (any nonzero standard prediction versus a true zero response) for fully-indirect pairs. Crucially, this overstatement factor depends only on the observable direct-routing share $S_{od,0}$, not on the calibrated σ or the recovered $\hat{\theta}$: regardless of which row in Table C1 one adopts, the ratio of standard-gravity to triadic elasticities is $1/S_{od,0}$ for every O-D pair.

To aggregate this attenuation into a single network-level number, we compute the value-weighted expected elasticity across all 128 O-D pairs: $\mathbb{E}_v[\varepsilon_{od}] = -\hat{\theta} \cdot \bar{S}^v_{od,0}$, where $\bar{S}^v_{od,0} = 0.154$ is the trade-value-weighted mean direct share. This gives $\mathbb{E}_v[\varepsilon_{od}] \approx -1.83$, compared with the standard-gravity prediction of −11.87 for every pair. This figure is specific to our baseline calibration $\sigma = 4$; Table C1 shows the overstatement factor ranges from 3.2× at $\sigma = 3$ to 13× at $\sigma = 6$. However, the pair-level overstatement ratio $1/S_{od,0}$ is invariant to σ: it depends only on the observed direct-

routing share, and the network-level ratio $1/\bar{S}^{v}_{od,0} = 1/0.154 \approx 6.5$ is a σ-free sufficient statistic for the aggregate overstatement. Concretely, a uniform 10% increase in bilateral iceberg costs across all O-D routes in this network would reduce aggregate trade value by approximately 18% under the triadic model, versus 69% under standard gravity—an overstatement of **6.5**×.[19] These numbers should be interpreted as specific to this trading house's network; the mechanism and the direction of attenuation are expected to generalize, but the exact magnitude depends on the share of indirect routing in the relevant network. These calculations hold destination wages and demand fixed (partial equilibrium); general-equilibrium wage adjustments would partially offset the direct trade-flow responses. The policy implication follows directly: trade policy analyses calibrated on aggregate bilateral gravity elasticities will systematically over-predict trade-flow losses from tariff increases in sectors with deep intermediation. The attenuation is strongest precisely for the pairs that are most fully intermediated—which are also, typically, the most geographically remote and cost-sensitive pairs.

5.2 Russia as Intermediary Hub: Substitution Patterns 2018–2022

The second exercise documents hub reorganization following Russia's decline as a petrochemical intermediary. Russia served as a petrochemical intermediary in our data beginning in 2019, with a pronounced peak in 2020–2021, followed by a sharp decline through 2022. Table 8 reports Russia's role as hub by year.

Table 8: Russia as intermediary: transaction count and trade value, 2019–2022

Year	Transactions (Russia hub)	Trade Value (Russia hub)
2019	12	45.9
2020	72	6599.7
2021	48	22217.5
2022	24	5048.0

Notes: Russia identified as intermediary country (`iso_m = RUS`).
Trade value is expressed in thousands of U.S. dollars.

[19] Using the equal-weight mean $\bar{S}_{od,0} = 0.124$ instead gives an overstatement factor of 8.0×. The value-weighted figure is the relevant benchmark for aggregate welfare calculations, since it weights pairs by their contribution to total trade. These calculations apply the first-order elasticity to a finite 10% shock. For fully-indirect pairs ($S_{od,0} = 0$), the prediction of zero response is exact for any shock size. For partially-direct pairs, a finite shock shifts $S_{od,0}$ itself; the first-order formula provides an approximation, but the direction of attenuation (triadic < standard gravity) is preserved at any shock size.

The timing is consistent with geopolitical disruption: Russia's intermediary role peaked in transaction count at 72 shipments in 2020 (USD 6.6M) and in trade value in 2021 (USD 22.2M), before collapsing to 24 transactions (USD 5.0M) in 2022—a decline of two-thirds in count and 77% in value from the 2021 peak. Table 9 shows how the broader hub network evolved over the same period, tracking the top-8 intermediary countries by total volume.

Table 9: Hub substitution: top intermediary countries by transaction count, 2018–2022

Intermediary	2018	2019	2020	2021	2022	Δ 2020–22 (%)
SGP	2724	756	432	72	132	-69.4%
VNM	516	564	264	552	312	18.2%
IND	648	468	72	324	12	-83.3%
CHN	0	0	228	348	324	42.1%
KOR	300	228	192	72	108	-43.7%
ECU	252	300	192	36	0	n/a
DEU	0	12	60	84	72	20.0%
JPN	84	60	24	24	24	0.0%

Notes: Transaction counts by intermediary country (`iso_m`) and year. Top 8 intermediaries by total 2018–2022 volume shown; Russia (RUS) is excluded as it does not rank in the top 8 by aggregate volume over this period; see Table 8 for Russia's year-by-year trajectory. Δ 2020–22 column shows % change in transaction count from 2020 (peak Russia year) to 2022. Positive values indicate growing hub role; negative values indicate decline.

The pattern is consistent with the model's prediction that hub disruption triggers reorganization along the intermediary network. China's role as hub grew substantially (0 transactions in 2018–2019, rising to 324 by 2022, +42% from 2020 to 2022), while Vietnam maintained a stable and growing intermediary role (+18% from 2020 to 2022). Singapore, the largest hub, declined sharply (−69% from 2020 to 2022)—consistent with a broader reorientation away from Singapore-routed petrochemical trade toward mainland Asia hubs during this period. These dynamics illustrate how the triadic routing mechanism operates at the network level: disruption at one hub node triggers substitution toward adjacent, lower-cost hubs rather than a proportional reduction in overall trade volume. The reorientation is also consistent with rules-of-origin effects: as sanctions altered the compliance and regulatory landscape for Russia-originating transactions, routing through jurisdictions with lower compliance friction became relatively more attractive (Conconi et al., 2018).

Upper bound on the Russia compliance-cost increase. The Russia episode is structurally distinct from the bilateral spoke-cost counterfactual in Section 5.1: it represents a *hub-cost shock*—

an increase in the compliance-and-logistics wedge $\hat{\gamma}_{\text{RUS}}$ at the intermediary node rather than a change in spoke distances. Under CES demand for hub services, indirect trade through hub m is proportional to $(\hat{\gamma}_m \cdot d_{mo} \cdot d_{md})^{1-\sigma}$. Attributing the entire observed volume decline to the hub-cost channel yields an upper bound on the implied increase in $\hat{\gamma}_{\text{RUS}}$, since part of the decline may also reflect demand-side factors such as sanctions affecting Russian buyers or broader recessionary effects. Specifically,

$$\frac{\hat{\gamma}_{\text{RUS,new}}}{\hat{\gamma}_{\text{RUS,old}}} = \left(\frac{X_{2022}}{X_{\text{peak}}}\right)^{1/(1-\sigma)}, \tag{7}$$

where $\sigma = 4$. The observed declines—66.7% in transaction count, 77.3% in trade value from the respective peaks—imply upper-bound $\hat{\gamma}$ ratios of 1.442 (count-based) and 1.639 (value-based), corresponding to compliance-cost increases of at most 44.2% and 63.9% respectively. Expressed as a distance equivalent, these are equivalent to adding 2923–4226 km to the hub-to-origin leg ($d_{mo} \approx 6{,}614$ km for the Korea–Russia route). Table 10 summarizes these calculations. Notably, 1% of Russia-hub volume by value was classified as destination-serving (buyers located in Russia), rather than transit routing through Russia to third destinations; this confirms that the Russia disruption affected the hub-as-destination function rather than the transit-routing function, which is consistent with the compliance-cost channel and the model's hub-demand formula.

Table 10: Russia hub-cost shock: observed declines and implied compliance-cost increase

	Transaction count	Trade value
Panel A: Observed Russia-hub activity		
Peak year (2020/2021)	72	22217.0
2022	24	5048.0
Decline from peak	66.7%	77.3%
Panel B: Implied compliance-cost increase ($\hat{\gamma}_{RUS,new}/\hat{\gamma}_{RUS,old}$)		
Implied $\hat{\gamma}$ ratio	1.442 (+44.2%)	1.639 (+63.9%)
Distance equivalent (added km)	2923 km	4226 km

Notes: Russia-hub transactions aggregated across all O-D pairs (`iso_m = RUS`). Peak years: 2020 for transaction count, 2021 for trade value. Trade value is expressed in thousands of U.S. dollars. Implied $\hat{\gamma}$ ratio from $\hat{\gamma}_{\text{RUS,new}}/\hat{\gamma}_{\text{RUS,old}} = (X_{2022}/X_{\text{peak}})^{1/(1-\sigma)}$, $\sigma = 4$, applied separately to count and value declines. Distance equivalent: km added to the hub-to-origin leg ($d_{mo} \approx 6{,}614$ km for the KOR–RUS route) that would generate the same proportional cost increase.

6 Conclusion

This paper documents rich triadic patterns in intermediated trade for petrochemical products and shows that these flows are systematically structured by the geography of origin–hub–destination legs. Using triadic gravity equations, we quantify how intermediated trade varies across the three legs of a triad. The evidence is consistent with a monotone, hierarchical selection into intermediaries: firms at nearby productivity cutoffs reallocate between the direct spoke and the cheapest adjacent indirect route when spoke costs shift. Because our data condition on observed intermediaries, the empirical design incorporates a single sufficient statistic, allowing leg coefficients to recover intensive-margin elasticities.

Building on these estimates, we formulate a triadic framework that nests the bilateral benchmark and delivers clear comparative statics: bilateral trade elasticities with respect to spoke costs are strictly attenuated relative to the canonical Melitz elasticity $-\theta$, lying in $[-\theta, 0]$ in the bilateral case and in $[-\theta \cdot S^{\text{arc}}, 0]$ in the aggregate (arc) case. The gap is governed by two observables at the triad level—the adjacent-switch weight (ASW) and boundary sensitivity (BSI)—which summarize the scope for re-routing. When the indirect route is prohibitively expensive, the model yields standard Melitz elasticities. However, as the indirect option becomes cost-competitive, firms increasingly switch modes to bypass direct-route shocks; this re-routing insulates aggregate trade volume, effectively driving the observed bilateral elasticity toward zero. Exploiting model-based selection controls, we recover a high structural elasticity parameter that indicates aggressive switching into intermediary trade in response to trade frictions. This mechanism explains why our triadic elasticities are smaller in magnitude than bilateral gravity estimates, implying that models or policies calibrated on bilateral elasticities alone will overstate the response of trade flows to spoke-specific shocks.

These findings have practical implications. First, policy shocks aimed at a single leg (e.g., spoke-side tariffs, sanctions, or logistics frictions) diffuse through re-routing margins, so their measured impact depends on the local tightness of the $0 \leftrightarrow 1$ boundary (direct vs. cheapest indirect) rather than on spoke costs alone. Second, hub-side wedges and multimodal connectivity shape both intensive responses and selection, so infrastructure and regulatory policy at intermediaries can materially buffer or amplify the shocks at the spokes. Third, identification in selected samples requires

explicit control for the moving boundary; without it, bilateral gravity designs confound intensive and extensive margins, biasing elasticities away from their true values. While our curvature-based sufficient statistic is parsimonious, relaxing the curvature constraint would require multiple switch controls—one per moving boundary—which we leave for future work, along with extensions to multi-product firms and time-varying hub wedges.

The triadic perspective clarifies when and why trade responds less than the bilateral benchmark: selection into intermediaries is not noise but the mechanism that attenuates elasticities. Recognizing and estimating this mechanism enhances counterfactual analysis of trade policy in intermediated networks, yielding more realistic predictions for the incidence and pass-through of shocks in global petrochemical supply chains.

References

Abel-Koch, J. (2013). Who uses intermediaries in international trade? evidence from firm-level survey data. *The World Economy*, 36(8):1041–1064.

Ahn, J., Khandelwal, A. K., and Wei, S.-J. (2011). The role of intermediaries in facilitating trade. *Journal of International Economics*, 84(1):73–85.

Akerman, A. (2018). A theory on the role of wholesalers in international trade based on economies of scope. *Canadian Journal of Economics/Revue canadienne d'économique*, 51(1):156–185.

Anderson, J. E. and van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, 93(1):170–192.

Antras, P. and Chor, D. (2013). Organizing the global value chain. *Econometrica*, 81(6):2127–2204.

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

Bernard, A. B., Blanchard, E. J., Van Beveren, I., and Vandenbussche, H. (2019). Carry-along trade. *The Review of Economic Studies*, 86(2):526–563.

Bernard, A. B., Grazzi, M., and Tomasi, C. (2015). Intermediaries in international trade: Products and destinations. *Review of Economics and Statistics*, 97(4):916–920.

Bernard, A. B., Jensen, J. B., Redding, S. J., and Schott, P. K. (2010). Intermediation in international trade: Wholesalers and retailers in US trade. *American Economic Review*, 100:408–413.

Blum, B. S., Claro, S., and Horstmann, I. (2010). Facts and figures on intermediated trade. *American Economic Review*, 100(2):419–423.

Broda, C. and Weinstein, D. E. (2006). Globalization and the gains from variety. *Quarterly Journal of Economics*, 121(2):541–585.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427.

Cameron, A. C. and Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372.

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *American Economic Review*, 98(4):1707–1721.

Conconi, P., Garcia-Santana, M., Puccio, L., and Venturini, R. (2018). From final goods to inputs: The protectionist effect of rules of origin. *American Economic Review*, 108(8):2335–2365.

Crozet, M., Lalanne, G., and Poncet, S. (2013). Wholesalers in international trade. *European Economic Review*, 58:1–17.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5):1741–1779.

Eaton, J., Kortum, S., and Kramarz, F. (2011). An anatomy of international trade: Evidence from french firms. *Econometrica*, 79(5):1453–1498.

Fally, T. and Hillberry, R. (2018). A Coasian model of international production chains. *Journal of International Economics*, 114:299–315.

Felbermayr, G. and Jung, B. (2011). Trade intermediation and the organization of exporters. *Review of International Economics*, 19(4):634–648.

Freund, C. and Pierola, M. D. (2015). Export superstars. *Review of Economics and Statistics*, 97(5):1023–1032.

Ganapati, S. (2025). The modern wholesaler: Global sourcing, domestic distribution, and scale economies. *American Economic Journal: Microeconomics*, 17(1):1–40.

Head, K. and Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. *Handbook of International Economics*, 4:131–195.

Helpman, E., Melitz, M. J., and Yeaple, S. R. (2004). Export versus FDI with heterogeneous firms. *American Economic Review*, 94(1):300–316.

Irarázabal, A., Moxnes, A., and Opromolla, L. D. (2013). The margins of multinational production and the role of intrafirm trade. *Journal of Political Economy*, 121(1):74–130.

Mayer, T. and Zignago, S. (2011). Notes on CEPII's distances measures: The geodist database. Working Papers 2011-25, CEPII.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics*, 48(1):7–35.

Silva, J. S. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658.

Simonovska, I. and Waugh, M. E. (2014). The elasticity of trade: Estimates and evidence. *Journal of International Economics*, 92(1):34–50.

Appendix

A Additional Tables and Figures

A.1 Product-level descriptive statistics and match patterns

Table A1 lists all 23 petrochemical products in the sample with average trade volumes, sales values, intermediary profits, and buyer/seller counts over 2018–2022.

Table A1: Descriptive statistics of intermediary trade across products

Product	Quantity	Sales value	Intermediary profit	No. buyers	No. sellers
AVGAS	6818.52	3831.14	39.32	37	14
Aromatics	86.96	113.58	0.86	10	9
Asphalt	16737.96	6278.12	80.29	27	6
Bunkering abroad	4173.74	1765.82	-5.17	9	9
Bunkering domestic	4984.73	2500.07	14.53	76	15
Butyl rubber	596.44	1134.54	12.23	3	1
Complex fertilizer	8304.03	3669.40	25.13	8	3
Fertilizer materials	6399.85	1264.85	19.13	6	4
Fuel oil	50211.43	22755.82	140.95	11	14
Gas oil	10052.69	4296.24	8.97	6	3
Jet oil	817.63	336.33	7.04	2	2
MOGAS	7604.49	4489.21	66.14	3	3
Nitrogen fertilizer	507.92	173.64	4.75	25	23
Olefins	16572.55	10679.25	17.61	6	3
Other petroleum products	8337.54	2484.61	10.59	45	6
Other polymer products	59.20	92.72	1.73	19	6
PE	346.28	407.71	5.04	76	20
PET	3191.96	3551.75	43.61	19	7
PP	285.08	295.35	2.70	37	11
Phosphatic fertilizer	2115.04	742.28	16.95	4	6
Plasticizer	172.03	212.00	3.60	20	3
Solvents	98.97	100.54	3.03	5	4
Tarpaulin	36.03	106.42	1.11	18	2

Notes: This table reports the average trade volume, sales value, intermediary profit, and the total number of distinct buyers and sellers over 2018–2022. All monetary values are expressed in thousands of U.S. dollars, and quantities are expressed in thousands of metric tons.

A.2 Cross-border fragmentation and co-location patterns: quantities, values, and profits

The pattern in Tables A2 to A4 parallels that documented in Table 2 for the number of transactions. Across all outcomes–quantities, values, and intermediary profits–intermediation remains strongly shaped by geography. The dominant pattern is destination-centric: transactions in which intermediaries share the buyer's country ($d = m$) account for roughly half of total trade across all metrics and increase steadily over time. In contrast, seller-side intermediation ($m = o$) declines markedly, mirroring the shrinkage observed in Table 2. Fully fragmented triads remain common but gradually give way to buyer-intermediary co-location as intermediaries embed themselves in final markets.

Table A2: Annual patterns of intermediary transactions by co-location type: quantities

Year	Total	Fragmented	Partially Fragmented			Domestic
		($o \neq d$, $d \neq m$, $m \neq o$)	($o = d \neq m$)	($d = m \neq o$)	($m = o \neq d$)	($o = d = m$)
2018	2.037	0.809	0.167	0.508	0.391	0.163
		(0.397)	(0.082)	(0.249)	(0.192)	(0.080)
2019	1.227	0.317	0.015	0.287	0.018	0.590
		(0.258)	(0.013)	(0.234)	(0.014)	(0.481)
2020	0.945	0.256	0.086	0.268	0.028	0.308
		(0.270)	(0.091)	(0.283)	(0.030)	(0.326)
2021	1.475	0.628	0.112	0.606	0.000	0.129
		(0.426)	(0.076)	(0.411)	(0.000)	(0.087)
2022	1.164	0.330	0.210	0.545	0.021	0.057
		(0.283)	(0.181)	(0.468)	(0.018)	(0.049)
All	6.849	2.339	0.590	2.214	0.458	1.247
		(0.342)	(0.086)	(0.323)	(0.067)	(0.182)

Notes: Column "Total" reports the aggregate quantity shipped across all matched transactions (in millions of metric tons). Symbols o, d, and m denote the origin, destination, and intermediary countries, respectively. Column "Fragmented" corresponds to transactions in which all three parties are located in different countries, while column "Domestic" corresponds to fully domestic transactions. Columns under "Partially Fragmented" indicate cases in which two parties share a country while the third is located elsewhere. Numbers in parentheses represent shares of the yearly total.

Table A3: Annual patterns of intermediary transactions by co-location type: values

Year	Total	Fragmented	Partially Fragmented			Domestic
		$(o \neq d,\ d \neq m,\ m \neq o)$	$(o = d \neq m)$	$(d = m \neq o)$	$(m = o \neq d)$	$(o = d = m)$
2018	1054.290	471.759	83.483	311.975	90.472	96.602
		(0.447)	(0.079)	(0.296)	(0.086)	(0.092)
2019	554.325	193.030	9.769	185.734	8.805	156.986
		(0.348)	(0.018)	(0.335)	(0.016)	(0.283)
2020	330.275	87.748	31.980	107.894	12.607	90.046
		(0.266)	(0.097)	(0.327)	(0.038)	(0.273)
2021	761.098	304.846	63.869	345.475	0.000	46.907
		(0.401)	(0.084)	(0.454)	(0.000)	(0.062)
2022	861.335	268.639	142.932	401.355	16.901	31.508
		(0.312)	(0.166)	(0.466)	(0.020)	(0.037)
All	3561.323	1326.022	332.033	1352.434	128.785	422.049

Notes: Column "Total" reports the total value of all matched transactions (in millions of U.S. dollars). Symbols o, d, and m denote the origin, destination, and intermediary countries, respectively. Column "Fragmented" corresponds to transactions in which all three parties are located in different countries, while column "Domestic" corresponds to fully domestic transactions. Columns under "Partially Fragmented" indicate cases in which two parties share a country while the third is located elsewhere. Numbers in parentheses represent shares of the yearly total.

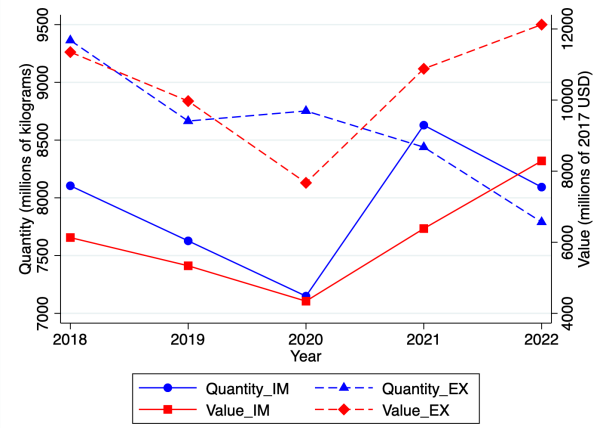
Table A4: Annual patterns of intermediary transactions by co-location type: profits

Year	Total	Fragmented	Partially Fragmented			Domestic
		$(o \neq d,\ d \neq m,\ m \neq o)$	$(o = d \neq m)$	$(d = m \neq o)$	$(m = o \neq d)$	$(o = d = m)$
2018	8.129	4.124	0.790	2.852	0.143	0.220
		(0.507)	(0.097)	(0.351)	(0.018)	(0.027)
2019	5.121	2.117	0.035	1.797	0.070	1.101
		(0.413)	(0.007)	(0.351)	(0.014)	(0.215)
2020	3.431	0.880	0.109	1.265	0.178	0.999
		(0.256)	(0.032)	(0.369)	(0.052)	(0.291)
2021	5.370	1.864	0.057	2.802	0.000	0.647
		(0.347)	(0.011)	(0.522)	(0.000)	(0.121)
2022	6.219	1.109	1.625	3.157	0.079	0.249
		(0.178)	(0.261)	(0.508)	(0.013)	(0.040)
All	28.270	10.094	2.616	11.873	0.470	3.217
		(0.357)	(0.093)	(0.420)	(0.017)	(0.114)

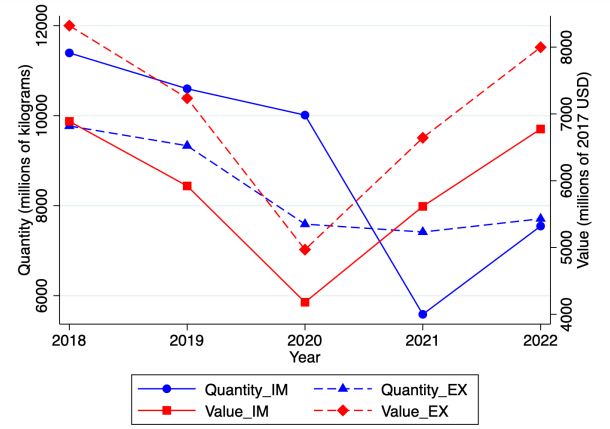
Notes: Column "Total" reports the total profit earned by intermediaries across all matched transactions (in millions of U.S. dollars). Symbols o, d, and m denote the origin, destination, and intermediary countries, respectively. Column "Fragmented" corresponds to transactions in which all three parties are located in different countries, while column "Domestic" corresponds to fully domestic transactions. Columns under "Partially Fragmented" indicate cases in which two parties share a country while the third is located elsewhere. Numbers in parentheses represent shares of the yearly total.

A.3 Evidence on aggregate trade trends and exit dynamics

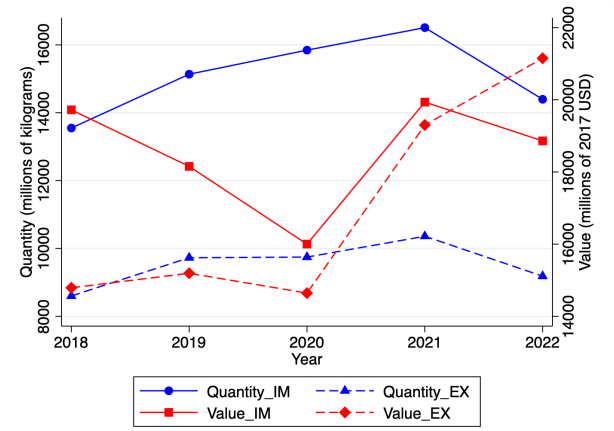
Figure A1 presents annual import and export patterns of petrochemical products for the three largest intermediary countries in our sample: South Korea, Singapore, and China, over 2018–2022. While intermediated matches decline over this period, these country-level trade flows show no comparable contraction. In both quantities and values, aggregate imports and exports remain relatively stable or even increase by 2022. This pattern rules out the possibility that the decline in intermediated transactions is driven by a shrinking petrochemical market.



(a) South Korea



(b) Singapore

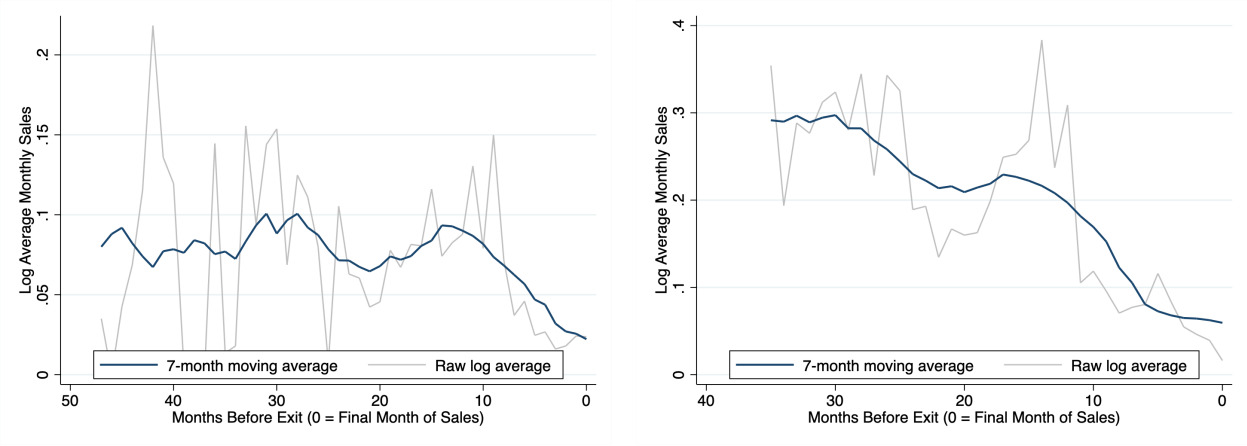


(c) China

Figure A1: Petrochemical imports and exports of top three intermediary countries

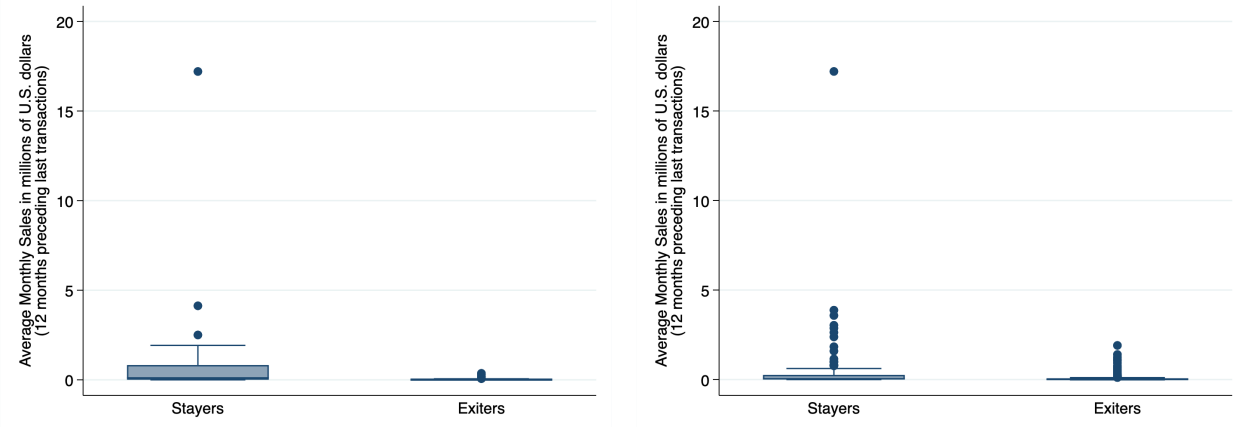
Figures A2 and A3 provide complementary evidence on the behavior of firms that exit intermediary relationships. Figure A2 plots the average log monthly sales of exiting sellers and buyers in the months leading up to their final transaction. In both cases, the smoothed series reveal a

gradual decline rather than a period of growth, weakening the interpretation that firms "graduate" out of intermediated trade due to expansion. Figure A3 compares the distribution of pre-exit sales between stayers and exiters. Exiters are systematically smaller and more concentrated near the lower end of the sales distribution, whereas stayers include a broader mass of larger firms. This stands in contrast to the well-documented pattern in which large or fast-growing firms substitute away from intermediaries and switch to direct trade.



(a) Sellers (b) Buyers

Figure A2: Average log sales for exiters



(a) Sellers (b) Buyers

Figure A3: Sales value comparison – stayers vs. exiters

A.4 Pairwise distance and trade outcomes

Figure A4 extends the visual evidence from Figure 3 by plotting two additional outcomes—shipped quantities (Panel A) and intermediary profits (Panel B)—against each of the three pairwise distances. The patterns are consistent: doubling any leg reduces shipped tonnage by 50–80% and profits by a comparable but somewhat smaller margin, suggesting that profit margins partly cushion the volume loss at longer distances.

Taken together, these plots reinforce the message of Figure 3: distance depresses both the probability of matching and the intensity of trade. The negative bivariate relationships stand in contrast to the positive coefficient on seller-intermediary distance that emerges once country-year fixed effects are introduced, indicating that the raw patterns conflate transport costs with intermediaries' strategic decisions to locate near major exporters.

A.5 Distribution of direct-trade shares

Table A5 reports the full distribution of $S_{od,0}$ across all 128 O-D pairs observed in the sample. The table underlies the attenuation calculations discussed in Section 5.1.

Table A5: Distribution of direct-trade share $S_{od,0}$ and implied triadic elasticities

Direct-share category	N O-D pairs	% of pairs	Mean $S_{od,0}$	Triadic ε_{od}
Fully indirect ($S = 0$)	109	84.5%	0.000	0.00
Low direct share ($0 < S < 0.25$)	2	1.6%	0.011	-0.13
Moderate direct share ($0.25 \leq S < 0.75$)	2	1.6%	0.278	-3.30
High direct share ($0.75 \leq S < 1$)	3	2.3%	0.818	-9.71
Fully direct ($S = 1$)	12	9.3%	1.000	-11.87
All O-D pairs	128	100.0%		

Notes: $S_{od,0}$ is measured directly from the raw transaction data as the share of total O-D trade value (summed across all years and products) transacted through the direct mode, defined as observations where the intermediary country equals the origin country (`iso_m = iso_o`). It is not backed out from the structural equation. Triadic $\varepsilon_{od} = -\hat{\theta} \cdot S_{od,0}$ is the first-order response of total O-D trade to a 1% bilateral cost increase (Proposition 3). Standard gravity predicts $-\hat{\theta} = -11.87$ for all pairs. Calibrated $\hat{\theta} = 11.87$, $\sigma = 4$. The 128 O-D pairs are the origin–destination routes observed in this trading house's transaction network; they do not represent all possible global petrochemical O-D pairs.

A.6 Robustness checks

The dyadic regression in Table 3 uses only transactions in which intermediaries are located in the same country as the seller. This restriction implies a strong selection effect: only seller-intermediary pairs that remain viable under domestic matching appear in the sample. For quantities, this selection dampens the distance elasticity toward zero, since sellers can already optimize their choice of local partner. For sales values, selection operates in the opposite direction: long-distance pairs that survive in the restricted sample tend to involve higher-value trades, generating a positive distance coefficient. To assess the sensitivity of these results, Table A6 reports an alternative dyadic specification that aggregates trade across all intermediary locations. Once the domestic restriction is removed, the conventional negative distance elasticity reappears for both quantities and values, confirming that the positive coefficient in Table 3 arises from the composition of surviving domestic-intermediary matches rather than a structural long-distance preference.

Table A6: PPML estimates of alternative specification for the dyadic gravity model

	(1) Quantity	(2) Sales Value	(3) Markup
$\log(\text{distance}_{od})$	-0.203*	-0.136*	0.086
	(0.121)	(0.070)	(0.078)
Origin-year FE	Yes	Yes	Yes
Destination-year FE	Yes	Yes	Yes
Product FE	Yes	Yes	Yes
Pseudo R-Squared	0.404	0.454	0.432
Log-likelihood	-1.75e+07	-5.12e+06	-3.46e+04
Observations	12052	11778	12037

Notes: The dependent variables are the intermediate trade outcomes (indicated by the column headers) between exporter i and importer j, aggregated across all intermediary locations. Bilateral controls are included but not reported. Robust standard errors (in parentheses) are clustered at the origin-destination level. Statistical significance is denoted by ***, **, and * at the 1%, 5%, and 10% levels, respectively.

A.7 Triadic gravity with selection-correction control

Table A7 extends the triadic specification from Table 3 by adding the Switch sufficient statistic derived in Section C.1. Columns (4)–(6) append the Switch term as an additional covariate, with the direct-distance coefficient constrained to $1 - \sigma = -3$ (i.e., $\sigma = 4$) via the offset method. The distance coefficients in the triadic legs (d_{om}, d_{md}) are qualitatively unchanged, confirming that the triadic gradients reflect intensive-margin variation rather than selection composition.

Note that $\widehat{\mathsf{Switch}}_{od}$ enters with a *negative* coefficient ($\hat{\kappa} \approx -2.75$) in this full-sample triadic regression, in contrast to the *positive* coefficient ($\hat{\kappa} \approx 2.96$) estimated in the direct-subsample structural regression (Table 5). This sign reversal is mechanically consistent with the model. In the full triadic sample, which pools both direct-routed ($\text{iso}_m = \text{iso}_o$) and indirect-routed ($\text{iso}_m \neq \text{iso}_o$) observations, a large Switch value—indicating a strong cost advantage for the direct lane—predicts *fewer* indirect transactions, not more. Because PPML is estimated over all routing modes jointly, the Switch term captures this selection margin as a negative coefficient on indirect volume. The structural interpretation ($\hat{\kappa} > 0$ reflecting selection into direct trade) applies only to the direct-mode subsample as specified in Equation (6). The full-sample negative coefficient is thus the expected empirical signature of the same mechanism, not a contradiction of it.

A.8 Gravity coefficient stability across the 2020 disruption

The steady decline in intermediated transactions between 2018 and 2022—driven in part by Russia's shrinking hub role—raises the concern that compositional changes in the active network could confound the gravity estimates. To test whether the triadic distance elasticities are stable across the pre- and post-2020 periods, Table A8 adds three interaction terms—each baseline distance leg interacted with a post-2020 indicator ($\mathbf{1}[\text{year} \geq 2021]$)—to the main triadic PPML specification.

The three interaction terms are jointly insignificant at the 5% level (F-test $p = 0.071$). The baseline distance coefficients in Column (1) and the post-2020 terms in Column (2) are consistent in sign and magnitude, supporting the interpretation that the triadic gravity patterns are stable across the network disruption period. The post-2020 interactions are therefore not included in the main specification.[20]

[20]The baseline coefficient for $\ln \text{dist}_{od}$ in Column (1) (-0.228, $N = 12{,}867$) differs slightly from the main triadic

Table A7: Triadic gravity with Switch control ($\sigma = 4$ offset)

	(1) Qty+Switch	(2) Revenue+Switch	(3) Markup+Switch
$\log(\text{distance}_{jk})$	-0.046	-0.042	0.297***
	(0.074)	(0.056)	(0.065)
$\log(\text{distance}_{ik})$	0.308***	0.076	0.121*
	(0.057)	(0.052)	(0.065)
$\widehat{\text{Switch}}_{od}$	-2.749***	-2.918***	-2.920***
	(0.090)	(0.051)	(0.041)
$\text{common language}_{ij}$	1.031	0.407	0.613*
	(0.736)	(0.560)	(0.329)
$\text{common language}_{jk}$	-0.905	-1.088**	-0.294
	(0.844)	(0.532)	(0.458)
$\text{common language}_{ik}$	-0.503	-0.023	-2.429***
	(0.741)	(0.690)	(0.495)
same colony_{ij}	-3.757***	-3.313**	-0.856**
	(1.329)	(1.353)	(0.369)
same colony_{jk}	-5.882***	-4.396***	-3.353*
	(1.705)	(1.500)	(1.718)
same colony_{ik}	7.701***	5.579***	1.408***
	(1.273)	(1.446)	(0.536)
contiguity_{ij}	-3.457**	-1.230	-0.512
	(1.532)	(1.576)	(0.485)
contiguity_{jk}	-3.533**	-2.542**	-1.966**
	(1.505)	(1.105)	(0.951)
contiguity_{ik}	-0.550	-1.829	-1.513**
	(1.328)	(1.387)	(0.690)
Origin-year FE	Yes	Yes	Yes
Destination-year FE	Yes	Yes	Yes
Intermediary FE	Yes	Yes	Yes
Item FE	Yes	Yes	Yes
Pseudo R-Square	0.483	0.494	0.464
Log-likelihood	-1.54e+07	-4.78e+09	-3.29e+04
N	12771	12559	12839

Notes: Standard errors in parentheses

Baseline triadic columns (without Switch control) appear in Table 3.

Switch control imposed with $\sigma = 4$ offset on $\log d_{ij}$.

SEs clustered at origin-intermediary-destination level.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A8: Gravity coefficient stability: post-2020 interactions (PPML, quantity)

	(1) Baseline	(2) Post-2020 stability
$\log d_{ij}$	-0.228***	-0.195**
	(0.083)	(0.084)
$\log d_{ik}$	0.286***	0.254**
	(0.082)	(0.099)
$\log d_{jk}$	0.142	0.155*
	(0.087)	(0.092)
$\log d_{ij} \times \mathbf{1}$[post-2020]		-0.555**
		(0.221)
$\log d_{ik} \times \mathbf{1}$[post-2020]		0.223
		(0.166)
$\log d_{jk} \times \mathbf{1}$[post-2020]		-0.301*
		(0.178)
Origin x Year FE	Yes	Yes
Dest x Year FE	Yes	Yes
Intermediary FE	Yes	Yes
Product FE	Yes	Yes
Pseudo R-Sq	0.445	0.448
Log-likelihood	-1.66e+07	-1.65e+07
Observations	12867	12867
F (interactions = 0)		.
p-value		0.071

Notes: Standard errors in parentheses

PPML estimation; dependent variable = quantity.

Post-2020 = 1 for years 2021–2022.

SE clustered at origin × destination × intermediary level.

F-test: joint significance of all three post-2020 interaction terms.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A.9 Endogeneity of office placement

The 25 overseas offices were established over the firm's operating history and could, in principle, be systematically located in markets with already-high trade intensity. If so, the intermediary-location fixed effects in the triadic specification would absorb a confounded blend of country characteristics and selected placement, potentially biasing the d_{om} coefficient. We report three complementary checks.

Balance test. Table A9 compares average 2018 (pre-sample baseline) log trade values between countries with and without a trading house office. Neither the seller-side ($p = 0.20$) nor the buyer-side ($p = 0.91$) differences are statistically significant, indicating that office countries were not systematically larger traders at the start of the sample.

Table A9: Balance test: 2018 trade intensity by office presence

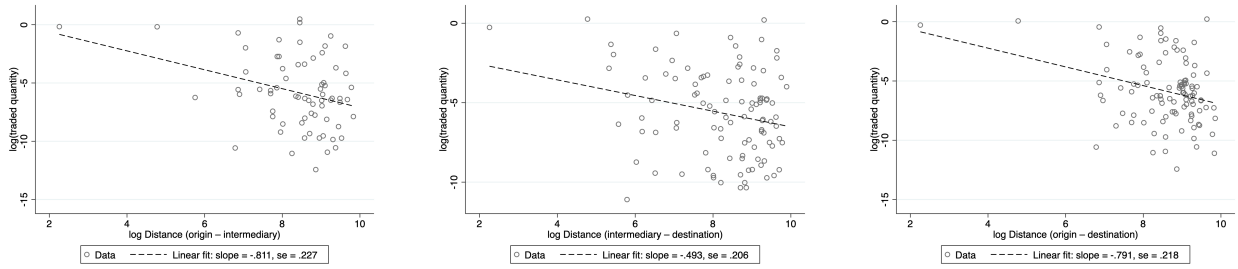
	(1)		
	No office	Has office	p-value
logval_o	9.00	13.47	0.200
logval_d	13.81	13.97	0.909
N	49		

Notes: Rows show mean log trade values in 2018 for countries without and with a trading house office. p-values from unequal-variance t-test.

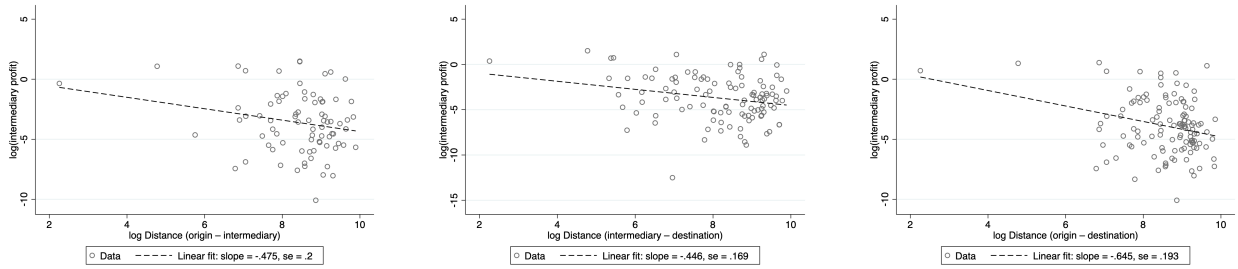
Logit of office presence on pre-period trade. Table A10 regresses an indicator for office presence on 2018 log trade values in a cross-country logit. Neither seller-side nor buyer-side trade volumes significantly predict office presence, further ruling out selection into high-trade markets.

First-year robustness. Table A11 repeats the triadic PPML regression restricting the sample to 2018 only—before any dynamic selection of office locations within the sample period could occur. In this smaller annual cross-section, the seller-hub distance coefficient (d_{om}) is negative (-0.255,

column in Table 3 (-0.221, $N = 12{,}771$). The difference reflects a broader sample: Table A8 includes the full 2018–2022 panel used for the interaction test, while the main specification in Table 3 drops observations with missing nationality controls. The qualitative patterns are identical.



A1. Quantity vs. d_{om} A2. Quantity vs. d_{md} A3. Quantity vs. d_{od}



B1. Profit vs. d_{om} B2. Profit vs. d_{md} B3. Profit vs. d_{od}

Figure A4: Triadic distance and trade outcomes: quantities (Panel A) and intermediary profits (Panel B). Each panel plots logged outcome against logged distance for one of the three legs. All outcomes decrease with distance on all legs.

Table A10: Logit: office presence on 2018 trade outcomes

	(1) Has trading house office
Log sales value (as seller, 2018)	0.177
	(0.124)
Log sales value (as buyer, 2018)	0.282
	(0.471)
Constant	-5.195
	(8.097)
Observations	12

Notes: Dependent variable = 1 if trading house has an office in the country. Robust standard errors. 2018 cross-section only.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

SE 0.167) and not statistically significant at conventional levels. The reduced precision is expected given the smaller sample and the high degree of within-year collinearity across distance legs when the panel dimension is removed. The buyer-hub coefficient (d_{md}), which drives the markup result, retains its positive sign. This robustness check suggests the triadic gradients are not an artifact of dynamic changes in office placement during the observation window.

Table A11: Triadic gravity: 2018 only (first-year robustness)

	(1) Quantity	(2) Sales Value	(3) Markup
log(distance$_{ij}$)	0.094***	0.078**	-0.002
	(0.028)	(0.039)	(0.020)
log(distance$_{jk}$)	-0.050	-0.059	0.363***
	(0.106)	(0.095)	(0.074)
log(distance$_{ik}$)	-0.255	-0.197	0.556**
	(0.167)	(0.148)	(0.258)
Origin FE	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes
Intermediary FE	Yes	Yes	Yes
Item FE	Yes	Yes	Yes
Pseudo R-Square	0.474	0.351	0.494
Log-likelihood	-3.75e+06	-1939.594	-1.20e+04
N	4784	4755	4842

Notes: Sample restricted to 2018 to test for dynamic selection of office locations within the sample period. The d_{ik} (seller-hub) coefficient

is negative and imprecisely estimated in this smaller annual cross-section; see main text for discussion. SEs clustered at origin-intermediary-destination level.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

B Model Details

B.1 Lemmas for Proposition 1

Lemma 2 (Single crossing)**.** *Under Assumption 1(a)–(b), fixed fees are strictly decreasing in k:*

$$f_{od,0} > f_{od,1} > \cdots > f_{od,M}.$$

Hence, for every adjacent pair $(k, k{+}1)$*,*

$$\left(\tau_{od,k+1}^{1-\sigma} - \tau_{od,k}^{1-\sigma}\right)\left(f_{od,k+1} - f_{od,k}\right) > 0,$$

so the two mode-specific profit lines cross exactly once (pairwise single crossing).

Proof. Assumption 1(a) gives $\tau_{od,0} < \cdots < \tau_{od,M}$, and since $1-\sigma < 0$, the map $\tau \mapsto \tau^{\,1-\sigma}$ is strictly decreasing, so $\tau_{od,k+1}^{1-\sigma} < \tau_{od,k}^{1-\sigma}$. By Assumption 1(b), $\frac{f_{od,k}}{\tau_{od,k}^{1-\sigma}} \geq \frac{f_{od,k+1}}{\tau_{od,k+1}^{1-\sigma}}$, so $f_{od,k+1} \leq f_{od,k}\,\frac{\tau_{od,k+1}^{1-\sigma}}{\tau_{od,k}^{1-\sigma}} < f_{od,k}$. Thus $f_{od,k+1} - f_{od,k} < 0$ while $\tau_{od,k+1}^{1-\sigma} - \tau_{od,k}^{1-\sigma} < 0$, and the product is positive. With linear profits in $\varphi^{\sigma-1}$, slopes strictly fall and intercepts strictly rise in k, implying single crossing. □

Lemma 3 (Zero-profit never binds)**.** *Under Assumption 1(a)–(b), for each interior mode* k*,*

$$\tilde{\varphi}_k^* \;\leq\; \tilde{\varphi}_{k,k+1} \quad \textit{(strict inequality generically).}$$

Hence, the lower edge of mode k*'s band is the* indifference *cutoff, not the zero-profit cutoff.*

Proof. Let $\tilde{\tau}_{od,k} \equiv \tau_{od,k}^{1-\sigma} > 0$, with $\tilde{\tau}_{od,k} > \tilde{\tau}_{od,k+1}$ by Assumption 1(a). Since $\sigma/A_d > 0$, $\tilde{\varphi}_k^* \leq \tilde{\varphi}_{k,k+1}$ is equivalent to

$$\frac{f_{od,k}}{\tilde{\tau}_{od,k}} \;\leq\; \frac{f_{od,k} - f_{od,k+1}}{\tilde{\tau}_{od,k} - \tilde{\tau}_{od,k+1}}.$$

Cross-multiplying by the positive denominators yields

$$\tilde{\tau}_{od,k}(f_{od,k} - f_{od,k+1}) \;\geq\; f_{od,k}(\tilde{\tau}_{od,k} - \tilde{\tau}_{od,k+1}) \iff \tilde{\tau}_{od,k} f_{od,k+1} \;\leq\; f_{od,k}\tilde{\tau}_{od,k+1}.$$

Dividing by $\tilde{\tau}_{od,k}\tilde{\tau}_{od,k+1} > 0$ gives

$$\frac{f_{od,k+1}}{\tilde{\tau}_{od,k+1}} \;\leq\; \frac{f_{od,k}}{\tilde{\tau}_{od,k}},$$

which is exactly Assumption 1(b). Strict inequality holds generically unless the normalized fee ratio is locally flat. □

While interior bands are bounded below by indifference (Lemma 3), zero–profit cutoffs can bind only at the outer edges: at the bottom of the lowest band and at the top band's upper tail.

B.2 Proof of trade elasticities

B.2.1 Case 1: identical iceberg factor across modes $\tau_{od,k} = \tau_{od}$.

When the composite iceberg term is the same for direct and indirect modes, every productivity threshold in eq. (4) scales linearly with τ_{od}.[21] Aggregating across the ladder recovers the canonical Melitz elasticity.

Lemma 4. *Suppose $\tau_{od,k} = \tau_{od}$ for all k and let $\tau_{od,k} = \Gamma_k \tau_{od}$ with $\Gamma_k > 0$ constant in τ_{od}. Then for any fixed $\{\Gamma_k, f_{od,k}\}$ and Pareto $\theta > \sigma - 1$,*

$$X_{od} = A_d T_o \sum_k \tau_{od,k}^{1-\sigma} Z_{od,k} \propto \tau_{od}^{-\theta} \qquad \Rightarrow \quad \varepsilon_{od} \equiv \frac{\partial \ln X_{od}}{\partial \ln \tau_{od}} = -\theta.$$

Proof. *From eq. (4), $\varphi_d^{(k \leftrightarrow k')} \propto \tau_{od}$ because $\tau_{od,k}^{1-\sigma} = \Gamma_k^{1-\sigma} \tau_{od}^{1-\sigma}$ so the common $\tau_{od}^{1-\sigma}$ factors out. Hence, each band term is $\tau_{od,k}^{1-\sigma} Z_{od,k} \propto (\Gamma_k^{1-\sigma} \tau_{od}^{1-\sigma}) \cdot (\tau_{od}^{\sigma-1-\theta}) = \Gamma_k^{1-\sigma} \tau_{od}^{-\theta}$. Summing over k preserves the $\tau_{od}^{-\theta}$ factor.* □

B.3 Proof of Proposition 3

Proof. **Step 0: Definitions.** Let $\nu \equiv \theta - (\sigma - 1) > 0$. Under Assumptions 1 and 2, only the boundary between $k = 0$ and $k = 1$ moves locally with τ_{od}. Define the origin-destination scale constant (incorporating the Pareto shape parameter):

$$\Omega_{od} \equiv A_d T_o. \tag{8}$$

Step 1: Route-level exports. Using the Pareto density $g(\varphi) = \theta \varphi^{-\theta-1}$, route-level exports are given by:

$$X_{od,k} = \Omega_{od} \cdot \tau_{od,k}^{1-\sigma} \left[(\varphi_{od,k}^{\text{low}})^{-\nu} - (\varphi_{od,k}^{\text{high}})^{-\nu} \right]. \tag{9}$$

[21] Examples of the "same line-haul" case, $\tau_{od,k} = \tau_{od}$, include (i) purely documentary intermediation (contracting, compliance, invoicing) and (ii) trade-finance/insurance overlays—escrow, credit insurance, letters of credit—layered on top of an unchanged physical shipment (no extra leg or handling). More generally, any proportional-scaling environment with $\tau_{od,k} = \Gamma_k \tau_{od}$, $\Gamma_k > 0$ constant in τ_{od} (and $f_{od,k}$ independent of τ_{od}), fits Lemma 4; examples include digital booking platforms that apply stable percentage fees and uniform policy/compliance surcharges.

For the direct mode ($k = 0$), the upper bound is $\varphi_{od,0}^{\text{high}} = \infty$. Thus:

$$X_{od,0} = \Omega_{od} \cdot \tau_{od,0}^{1-\sigma} (\varphi_{0,1})^{-\nu}, \quad \text{where } \varphi_{0,1} \equiv \varphi_{od}^{(0 \leftrightarrow 1)}. \tag{10}$$

For the adjacent indirect mode ($k = 1$), the bounds are the moving cutoff $\varphi_{0,1}$ and the next cutoff $\varphi_{1,2}$:

$$X_{od,1} = \Omega_{od} \cdot \tau_{od,1}^{1-\sigma} \left[(\varphi_{1,2})^{-\nu} - (\varphi_{0,1})^{-\nu} \right]. \tag{11}$$

Note that $\varphi_{1,2} \equiv \varphi_{od}^{(1 \leftrightarrow 2)}$ is determined by the intersection of modes $k = 1$ and $k = 2$, neither of which depends on $\tau_{od,0}$. Thus, for $k \geq 2$, both bounds are locally constant with respect to $\tau_{od,0}$, implying $\partial X_{od,k} / \partial \ln \tau_{od,0} = 0$.

Step 2: The moving boundary. The indifference condition between mode 0 and mode 1 is:

$$\frac{A_d}{\sigma} (\tau_{od,0}^{1-\sigma} - \tau_{od,1}^{1-\sigma}) (\varphi_{0,1})^{\sigma-1} = f_{od,0} - f_{od,1}. \tag{12}$$

Let $D_{od} \equiv \tau_{od,0}^{1-\sigma} - \tau_{od,1}^{1-\sigma}$. Taking logs:

$$(\sigma - 1) \ln \varphi_{0,1} = -\ln(D_{od}) + \text{constant}. \tag{13}$$

Differentiating with respect to $\ln \tau_{od,0}$:

$$\frac{\partial \ln \varphi_{0,1}}{\partial \ln \tau_{od,0}} = -\frac{1}{\sigma - 1} \frac{\partial \ln D_{od}}{\partial \ln \tau_{od,0}} = -\frac{1}{\sigma - 1} \frac{(1-\sigma) \tau_{od,0}^{1-\sigma}}{\tau_{od,0}^{1-\sigma} - \tau_{od,1}^{1-\sigma}} = \frac{\tau_{od,0}^{1-\sigma}}{\tau_{od,0}^{1-\sigma} - \tau_{od,1}^{1-\sigma}} \equiv \Psi_{od}. \tag{14}$$

Note that $\Psi_{od} > 1$ since $\tau_{od,0} < \tau_{od,1}$ implies $\tau_{od,0}^{1-\sigma} > \tau_{od,1}^{1-\sigma}$.

Step 3: Mode-specific elasticities. For the direct mode ($k = 0$), log-differentiating $X_{od,0}$:

$$\frac{\partial \ln X_{od,0}}{\partial \ln \tau_{od,0}} = \frac{\partial \ln(\tau_{od,0}^{1-\sigma})}{\partial \ln \tau_{od,0}} - \nu \frac{\partial \ln \varphi_{0,1}}{\partial \ln \tau_{od,0}} \tag{15}$$

$$\varepsilon^0 = (1 - \sigma) - \nu \Psi_{od}. \tag{16}$$

For the indirect mode ($k = 1$), $\tau_{od,1}$ is constant with respect to $\tau_{od,0}$, so only the lower bound of

the integral moves:

$$\frac{\partial \ln X_{od,1}}{\partial \ln \tau_{od,0}} = \frac{1}{X_{od,1}} \cdot \Omega_{od} \tau_{od,1}^{1-\sigma} \left[-(-\nu)(\varphi_{0,1})^{-\nu-1} \frac{\partial \varphi_{0,1}}{\partial \ln \tau_{od,0}} \cdot \varphi_{0,1} \right] \tag{17}$$

$$= \frac{\nu(\varphi_{0,1})^{-\nu}}{(\varphi_{1,2})^{-\nu} - (\varphi_{0,1})^{-\nu}} \frac{\partial \ln \varphi_{0,1}}{\partial \ln \tau_{od,0}}. \tag{18}$$

Using the definition of shares and Ψ_{od}:

$$\varepsilon^1 = \nu \Psi_{od} \left[\frac{(\varphi_{0,1})^{-\nu}}{(\varphi_{1,2})^{-\nu} - (\varphi_{0,1})^{-\nu}} \right]. \tag{19}$$

Step 4: Aggregate elasticity. The total bilateral elasticity is the share-weighted sum of mode elasticities:

$$\varepsilon_{od} = \sum_k S_{od,k} \varepsilon^k = S_{od,0} \varepsilon^0 + S_{od,1} \varepsilon^1. \tag{20}$$

Substituting eq. (16) and defining the adjacent-switch weight Λ_{od}:

$$\Lambda_{od} \equiv S_{od,1} \frac{(\varphi_{0,1})^{-\nu}}{(\varphi_{1,2})^{-\nu} - (\varphi_{0,1})^{-\nu}}, \tag{21}$$

we obtain:

$$\varepsilon_{od} = S_{od,0}[(1-\sigma) - \nu \Psi_{od}] + \nu \Psi_{od} \Lambda_{od} \tag{22}$$

$$= -(\sigma - 1) S_{od,0} - [\theta - (\sigma - 1)] \Psi_{od} (S_{od,0} - \Lambda_{od}). \tag{23}$$

This recovers eq. (5). □

B.4 Proof of Lemma 1

Proof. Under Pareto productivity with shape θ and minimum $\varphi_{\min}$, aggregate trade in each mode k is

$$X_{od,k} = \Omega_{od} \, \tau_{od,k}^{1-\sigma} \int_{\varphi_{k,k+1}}^{\varphi_{k-1,k}} \varphi^{\sigma-1-\theta-1} \, d\varphi,$$

where $\Omega_{od} \equiv A_d T_o$ is a common scale term, $\varphi_{0,\infty} \equiv \infty$ (upper bound for $k = 0$), and $\nu \equiv \theta - (\sigma - 1) > 0$. Evaluating the integral:

$$X_{od,0} = \frac{\Omega_{od}}{\nu}\, \tau_{od,0}^{1-\sigma} (\varphi_{0,1})^{-\nu}, \qquad X_{od,1} = \frac{\Omega_{od}}{\nu}\, \tau_{od,1}^{1-\sigma} \left[(\varphi_{1,2})^{-\nu} - (\varphi_{0,1})^{-\nu} \right].$$

Recalling that $\Lambda_{od} = S_{od,1} \cdot \frac{(\varphi_{0,1})^{-\nu}}{(\varphi_{1,2})^{-\nu} - (\varphi_{0,1})^{-\nu}}$, substituting gives

$$\Lambda_{od} = \frac{\Omega_{od}}{\nu\, X_{\text{total}}}\, \tau_{od,1}^{1-\sigma} (\varphi_{0,1})^{-\nu}, \qquad S_{od,0} = \frac{\Omega_{od}}{\nu\, X_{\text{total}}}\, \tau_{od,0}^{1-\sigma} (\varphi_{0,1})^{-\nu}.$$

Hence $S_{od,0}/\Lambda_{od} = (\tau_{od,0}/\tau_{od,1})^{1-\sigma}$. Since $\tau_{od,0} < \tau_{od,1}$ (direct route is cheaper) and $1 - \sigma < 0$, we have $(\tau_{od,0}/\tau_{od,1})^{1-\sigma} > 1$, so $S_{od,0} > \Lambda_{od}$. □

B.5 Proof of Corollary 1

Proof. From the proof of Proposition 3:

$$\varepsilon_{od} = -(\sigma - 1) S_{od,0} \; - \; [\theta - (\sigma - 1)]\, \Psi_{od}\, (S_{od,0} - \Lambda_{od}).$$

Lower bound ($\varepsilon_{od} \geq -\theta$). From the expressions in Section B.4 (proven above), we have the key identity $\Psi_{od}(S_{od,0} - \Lambda_{od}) = S_{od,0}$. To verify: $S_{od,0} - \Lambda_{od} = \frac{\Omega_{od}/\nu}{X_{\text{total}}} (\varphi_{0,1})^{-\nu} [\tau_{od,0}^{1-\sigma} - \tau_{od,1}^{1-\sigma}]$, and $\Psi_{od} = \tau_{od,0}^{1-\sigma} / [\tau_{od,0}^{1-\sigma} - \tau_{od,1}^{1-\sigma}]$, so their product equals $\frac{\Omega_{od}/\nu}{X_{\text{total}}} \tau_{od,0}^{1-\sigma} (\varphi_{0,1})^{-\nu} = S_{od,0}$. Substituting into the elasticity formula:

$$\varepsilon_{od} = -(\sigma - 1) S_{od,0} - [\theta - (\sigma - 1)] S_{od,0} = -\theta\, S_{od,0} \geq -\theta,$$

with equality as $S_{od,0} \to 1$.

Upper bound ($\varepsilon_{od} \leq 0$). Since $\varepsilon_{od} = -\theta S_{od,0}$ and $S_{od,0} \geq 0$, we have $\varepsilon_{od} \leq 0$.

Limits. As $\tau_{od,1} \to \infty$, $D_{od} \to \tau_{od,0}^{1-\sigma}$, the cutoff $\varphi_{0,1}$ is driven by $f_{od,0}/D_{od} \to 0$, so $S_{od,0} \to 1$ and $\varepsilon_{od} \to -\theta$. As $\tau_{od,0} \to \tau_{od,1}$, $D_{od} \to 0$, so $\varphi_{0,1} \to \infty$ and $S_{od,0} \propto \varphi_{0,1}^{-\nu} \to 0$. Since $\varepsilon_{od} = -\theta S_{od,0}$, we get $\varepsilon_{od} \to 0$. □

B.6 Multi-origin arc elasticity

Lemma 5 (Arc elasticity baseline under path–invariant iceberg costs)**.** *Suppose Case 1 holds: for every origin o and hub k, the composite iceberg is path–invariant, $\tau_{od,k} = \tau_{od}$, so that for each feasible mode k on $o \to d$ one can write $\tau_{od,k} = \Gamma_k \tau_{od}$ with $\Gamma_k > 0$ independent of τ_{od}. Let $X_{\cdot d}^{\mathrm{arc}(od)}$ denote the total value of d's imports that traverse the physical leg (o, d) (including direct $o \to d$ and any third–country traffic routed via o), and define the arc throughput share*

$$\rho_{od}^{\mathrm{arc}} \;\equiv\; \frac{X_{od}}{X_{\cdot d}^{\mathrm{arc}(od)}} \in [0, 1].$$

Then the arc elasticity with respect to τ_{od} equals

$$\varepsilon_{\cdot d}^{\mathrm{arc}(od)} \;\equiv\; \frac{\partial \ln X_{\cdot d}^{\mathrm{arc}(od)}}{\partial \ln \tau_{od}} \;=\; -\,\theta\, \rho_{od}^{\mathrm{arc}}.$$

Proof. Under path–invariance, any $u \to d$ flow with $u \neq o$ that is routed via o has composite cost independent of τ_{od}, so $\partial X_{ud}^{\mathrm{via}\ o} / \partial \ln \tau_{od} = 0$. By Lemma 4, $X_{od} \propto \tau_{od}^{-\theta}$, hence $\partial \ln X_{od} / \partial \ln \tau_{od} = -\theta$. Therefore

$$\frac{\partial X_{\cdot d}^{\mathrm{arc}(od)}}{\partial \ln \tau_{od}} = \frac{\partial X_{od}}{\partial \ln \tau_{od}} = -\theta\, X_{od},$$

and dividing by $X_{\cdot d}^{\mathrm{arc}(od)}$ yields $-\theta\, \rho_{od}^{\mathrm{arc}}$. □

Proposition 4 (Aggregate import elasticity to a spoke shock with hub spillovers)**.** *Work under Case 2 ($\tau_{ud,k} = \tau_{uk}\tau_{kd}$) and Assumptions 1 and 2. Fix a destination d and a physical leg (o, d). Let total imports into d be $X_{\cdot d} \equiv \sum_u X_{ud}$ and origin weights be $W_{ud} \equiv X_{ud}/X_{\cdot d}$.*

For the physical link (o, d), define the arc share in d's imports *as*

$$S_{\cdot d}^{\mathrm{arc}(od)} \;\equiv\; W_{od}\, S_{od,0} \;+\; \sum_{u \neq o} W_{ud}\, S_{ud}^{(o)},$$

where $S_{od,0}$ is the direct share from origin o, and $S_{ud}^{(o)}$ is the share of upstream origin u's exports that route via hub o.

For the local bilateral $o \to d$, let Ψ_{od} and Λ_{od} be defined as in Proposition 3. For each upstream origin $u \neq o$ routing through o, let $\tau_{ud,o} \equiv \gamma_o \tau_{uo} \tau_{od}$ be the composite cost. Let $j(u)$ denote the

adjacent competitor mode for origin u (either direct or another hub). Define the upstream boundary sensitivity as:

$$\Psi_{u;od} \;\equiv\; \frac{\tau_{ud,o}^{1-\sigma}}{\tau_{ud,o}^{1-\sigma} - \tau_{u,j(u)}^{1-\sigma}}.$$

Let $\tilde{\varphi}_u^{(o\leftrightarrow j)}$ be the moving indifference cutoff for origin u between hub o and option j, and let $\tilde{\varphi}_u^{j\uparrow}$ be the fixed upper bound of the adjacent band. With $\nu \equiv \theta - (\sigma - 1) > 0$, define the upstream switch weight:

$$\Lambda_{u;od} \;\equiv\; S_{ud}^{j(u)} \,\frac{(\tilde{\varphi}_u^{(o\leftrightarrow j)})^{-\nu}}{(\tilde{\varphi}_u^{j\uparrow})^{-\nu} - (\tilde{\varphi}_u^{(o\leftrightarrow j)})^{-\nu}} \;\in\; [\,0,\; 1 - S_{ud}^{(o)}\,].$$

Then the elasticity of total imports into d with respect to the spoke cost τ_{od} is

$$\varepsilon_{\cdot d}^{(od)} = -(\sigma-1)\, S_{\cdot d}^{\mathrm{arc}(od)} - \big[\theta-(\sigma-1)\big] \left[W_{od}\, \Psi_{od}\, \big(S_{od,0} - \Lambda_{od}\big) + \sum_{u\neq o} W_{ud}\, \Psi_{u;od}\, \big(S_{ud}^{(o)} - \Lambda_{u;od}\big) \right]. \quad (24)$$

Moreover,

$$0 \;\geq\; \varepsilon_{\cdot d}^{(od)} \;\geq\; -\,\theta\, S_{\cdot d}^{\mathrm{arc}(od)}. \quad (25)$$

Proof. Write $X_{\cdot d} = \sum_u X_{ud}$, so $\frac{\partial \ln X_{\cdot d}}{\partial \ln \tau_{od}} = \sum_u W_{ud}\, \frac{\partial \ln X_{ud}}{\partial \ln \tau_{od}}$.

(i) Local origin $u = o$. This recovers Proposition 3:

$$\frac{\partial \ln X_{od}}{\partial \ln \tau_{od}} = -(\sigma-1)\, S_{od,0} - \big[\theta - (\sigma-1)\big]\, \Psi_{od}\, \big(S_{od,0} - \Lambda_{od}\big).$$

(ii) Upstream origins $u \neq o$. Only the specific mode passing through o loads the leg (o, d). Since $\tau_{ud,o} = \Gamma \tau_{od}$, we have $\frac{\partial \ln \tau_{ud,o}}{\partial \ln \tau_{od}} = 1$. Under the curvature ladder, exactly the boundary between mode o and its adjacent competitor $j(u)$ moves. Repeating the two-band calculation from Proposition 3:

$$\frac{\partial \ln X_{ud}}{\partial \ln \tau_{od}} = -(\sigma-1)\, S_{ud}^{(o)} - \big[\theta - (\sigma-1)\big]\, \Psi_{u;od}\, \big(S_{ud}^{(o)} - \Lambda_{u;od}\big).$$

(iii) Aggregate. Weight (i) and (ii) by W_{ud} and sum. The first term sums to $-(\sigma-1) S_{\cdot d}^{\mathrm{arc}(od)}$. The second term aggregates the selection offsets. The bounds follow because for every origin, the selection term attenuates the response, ensuring the elasticity is strictly between 0 and the Melitz benchmark applied to the active arc share. □

C Selection into Intermediaries

C.1 Derivation of the Sufficient Statistic

We consider the firm's choice between direct exporting ($k = 0$) and the set of available indirect routes. Under the monotone ladder structure (Assumption 2), firms sort based on their productivity φ. A firm chooses the direct mode if its profit from direct exporting exceeds the profit from the best indirect option.

Let $\tau_{od,0}$ denote the variable cost of direct trade, and let $\tau_{od,1} \equiv \min_m(\gamma_m \tau_{om} \tau_{md})$ denote the composite variable cost of the most efficient indirect route. The indifference cutoff $\bar{\varphi}_{0,1}$ between the direct mode and the best indirect mode is defined by the condition $\pi_{od,0}(\bar{\varphi}_{0,1}) = \pi_{od,1}(\bar{\varphi}_{0,1})$. Using the standard Melitz profit function $\pi(\varphi) \propto (\tau/\varphi)^{1-\sigma} - f$, this equality implies:

$$\frac{\tau_{od,0}^{1-\sigma}}{\bar{\varphi}_{0,1}^{\sigma-1}} - f_{od,0} = \frac{\tau_{od,1}^{1-\sigma}}{\bar{\varphi}_{0,1}^{\sigma-1}} - f_{od,1}. \tag{26}$$

Rearranging for the cutoff productivity yields:

$$\bar{\varphi}_{0,1}^{\sigma-1} = \frac{f_{od,0} - f_{od,1}}{\tau_{od,0}^{1-\sigma} - \tau_{od,1}^{1-\sigma}}. \tag{27}$$

Aggregate trade flows in the direct mode ($X_{od,0}$) are determined by integrating firm revenues over the range of firms that select into direct exporting, $[\bar{\varphi}_{0,1}, \infty)$. Assuming Pareto-distributed productivity with shape parameter θ, the aggregate volume is proportional to $\bar{\varphi}_{0,1}^{-\nu}$ (where $\nu \equiv \theta - (\sigma - 1)$). Since $\bar{\varphi}_{0,1}^{\sigma-1} \propto D_{od}^{-1}$ where $D_{od} = \tau_{od,0}^{1-\sigma} - \tau_{od,1}^{1-\sigma}$, we have $\bar{\varphi}_{0,1}^{-\nu} \propto D_{od}^{\nu/(\sigma-1)} = D_{od}^{\kappa}$. Taking logs, the selection term enters the gravity equation with a *positive* sign:

$$\ln X_{od,0} \propto + \left(\frac{\theta - (\sigma - 1)}{\sigma - 1} \right) \ln \left(\tau_{od,0}^{1-\sigma} - \tau_{od,1}^{1-\sigma} \right). \tag{28}$$

Intuitively, a larger D_{od} (direct route more advantageous) lowers the indifference cutoff, admits more direct exporters, and raises $X_{od,0}$. To operationalize this, we define the observable sufficient statistic:

$$\widehat{\mathsf{Switch}}_{od} \equiv \frac{1}{\sigma - 1} \ln \left(\tau_{od,0}^{1-\sigma} - \tau_{od,1}^{1-\sigma} \right). \tag{29}$$

This term captures the variation in the extensive margin boundary induced by the cost gap between the direct and indirect modes.

C.2 Iterative Estimation Procedure

Joint identification of σ and κ would require iterating between a direct-subsample regression (recovering σ from the distance coefficient) and an indirect-subsample regression (recovering hub wedges γ_m from mode shares) until convergence. In practice, $\ln \text{dist}_{od}$ and $\widehat{\mathsf{Switch}}_{od}$ are highly collinear in our sample: the Switch term is functionally a log-difference of distances, so $\hat{\sigma}$ drifts toward 1 in every iteration and the cost-gap term vanishes. Because joint identification is infeasible in these data, we adopt the calibration strategy described in Section 4.7: fix $\sigma = 4$ and estimate κ conditional on it. Section C.3 reports sensitivity to the calibrated σ.

C.3 Sensitivity of $\hat{\theta}$ to the imposed σ

Table C1 reports the estimated selection parameter $\hat{\kappa}$ and the implied Pareto shape $\hat{\theta} = \hat{\kappa}(\sigma - 1) + (\sigma - 1)$ across a wide range of the imposed CES elasticity σ. Panel A covers the differentiated-goods range $\sigma \in \{3, 4, 5, 6\}$ from Head and Mayer's meta-analysis; Panel B covers the petrochemical-appropriate range $\sigma \in \{8, 10\}$ motivated by Broda and Weinstein (2006)'s product-level estimates for HS chapters 27–29.[22]

Two patterns emerge. First, $\hat{\kappa}$ is remarkably stable throughout: the standard error is ≈ 0.102 in every row, and the point estimate rises smoothly with σ (as required by the normalization $\hat{\theta} = \hat{\kappa}(\sigma-1)+(\sigma-1)$). The direction of attenuation—elasticities strictly below the Melitz benchmark—is therefore robust to the entire range of plausible σ values. Second, the implied $\hat{\theta}$ is highly sensitive to the calibrated σ: it ranges from 5.9 at $\sigma = 3$ to 89.0 at $\sigma = 10$. The petrochemical-range estimates ($\hat{\theta} \approx 55$–89) are implausibly high relative to the literature benchmarks of $\hat{\theta} \approx 4$–8 (Simonovska and Waugh, 2014; Eaton et al., 2011), which provides additional support for using the differentiated-goods range $\sigma \in \{3, 4, 5, 6\}$ for this market. Our preferred $\sigma = 4$ estimate of $\hat{\theta} \approx 11.9$ is somewhat above the cross-country average, consistent with the relatively standardized nature of petrochemical products; higher $\hat{\theta}$ indicates lower firm heterogeneity and thus stronger selection on trade costs.

[22] We exclude $\sigma = 12$ because at that value the switch term becomes perfectly collinear with the origin–destination fixed effects in our sample, so the coefficient is not identified.

We caution that $\hat{\theta}$ should be read as a sufficient statistic for the selection-margin sensitivity rather than as an independently identified structural parameter.

A note on the confidence intervals: the 95% CI column applies the delta method *conditional on* the calibrated σ, i.e., $\hat{\theta} \pm 1.96 \times (\sigma - 1) \times \widehat{SE}(\hat{\kappa})$. These intervals therefore capture only sampling uncertainty in $\hat{\kappa}$ and not the additional uncertainty that would arise from treating σ as unknown. The appropriate measure of calibration uncertainty is the cross-row spread of $\hat{\theta}$, which is substantial: moving from $\sigma = 3$ to $\sigma = 6$ changes $\hat{\theta}$ by a factor of five. Readers should interpret the within-row CIs accordingly.

Table C1: Sensitivity of $\hat{\kappa}$ and $\hat{\theta}$ to the imposed σ

σ	$\hat{\kappa}$	(SE)	$\hat{\theta} = \hat{\kappa}(\sigma - 1) + (\sigma - 1)$	95% CI for $\hat{\theta}$ (given σ)
Panel A: Differentiated-goods range				
3	1.969	(0.102)	5.94	[5.54, 6.34]
4	2.958	(0.102)	11.87	[11.27, 12.47]
5	3.947	(0.102)	19.79	[18.99, 20.59]
6	4.936	(0.102)	29.68	[28.68, 30.68]
Panel B: Petrochemical-appropriate range (Broda & Weinstein 2006)				
8	6.914	(0.102)	55.40	[54.00, 56.80]
10	8.892	(0.102)	89.03	[87.23, 90.84]

Notes: CI treats σ as calibrated and reflects sampling uncertainty in $\hat{\kappa}$ only, via the delta method: $\hat{\theta} \pm 1.96 \times (\sigma - 1) \times \widehat{SE}(\hat{\kappa})$. The spread of $\hat{\theta}$ *across rows* captures sensitivity to the choice of σ.

C.4 Seller-size robustness: ruling out size-sorting as the driver of d_{om}

The positive coefficient on d_{om} (seller-hub distance) in Table 3 could in principle reflect a size-sorting pattern: large sellers may concentrate in remote hubs, so the positive coefficient would capture seller size rather than a genuine geography effect. Table C2 tests this by adding log total

seller sales—computed as the sum of transaction value across all of a given exporter's shipments—as a direct control in the triadic PPML.

The d_{om} coefficient for quantity (0.319, SE 0.051, $p < 0.001$) is essentially unchanged relative to the baseline in Table 3. The seller-size control itself enters positively and significantly for quantity and sales value, confirming that larger exporters do transact more, but this additional variation does not absorb the hub-distance effect. We conclude that size-sorting is not the primary driver of the positive d_{om} coefficient, and the supply-side pooling mechanism in Section 4 remains the preferred interpretation.

Table C2: Triadic gravity with seller size control

	(1) Quantity	(2) Sales Value	(3) Markup
log(distance$_{ij}$)	-0.218***	-0.048	-0.066
	(0.084)	(0.085)	(0.061)
log(distance$_{ik}$)	0.319***	0.066	0.003
	(0.051)	(0.062)	(0.061)
log(distance$_{jk}$)	0.125*	-0.030	0.313***
	(0.070)	(0.059)	(0.052)
Log seller total sales (size proxy)	0.386***	0.459***	-0.051***
	(0.085)	(0.102)	(0.017)
Origin-year FE	Yes	Yes	Yes
Destination-year FE	Yes	Yes	Yes
Intermediary-year FE	Yes	Yes	Yes
Product FE	Yes	Yes	Yes
Pseudo R-Squared	0.455	0.487	0.457
Log-likelihood	-1.63e+07	-4.86e+09	-3.34e+04
Observations	12867	12643	12923

Notes: Standard errors in parentheses

Seller total sales computed as sum of transaction value across all transactions for each exporter firm.

If the positive d_{ik} (seller-hub) coefficient were driven by size sorting, it would shrink toward zero after controlling for log seller total sales.

SEs clustered at origin-intermediary-destination level.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$