RIIT 貿易障壁研究所

RIIT Working Paper Series

RWP-2026-005

Revisiting Trade Product-Type Classification: Substitution Elasticities, Ordered Alternatives, and a Comparison of the Rauch and IHY Typologies

Naohiko Ijiri
(RIIT, College of Economics, Nihon University)

Shuo Chen
(RIIT, Graduate School of Economics, Nihon University)

July 2026

Research Institute for Impediments to Trade

Tokyo, Japan

HP: https://riit.or.jp

Revisiting Trade Product-Type Classification: Substitution Elasticities, Ordered Alternatives, and a Comparison of the Rauch and IHY Typologies

Naohiko Ijiri* Shuo Chen†

July 2026

Abstract

Product-type classifications are widely used in the empirical international trade literature as proxies for differences in market organisation, search frictions, and substitutability across traded products. The most influential framework is the typology of Rauch (1999), which classifies products as organised-exchange, reference-priced, or differentiated. However, the Rauch classification is constructed at the SITC level and partly relies on qualitative criteria, raising concerns about transparency, reproducibility, and compatibility with modern HS-based trade data.

This paper provides a systematic empirical assessment of product-type classification systems by comparing the Rauch typology with the HS 6-digit-based IHY classification proposed by Ijiri et al. (2013). Both frameworks are motivated by the premise that differences in market organisation and standardisation translate into systematic differences in substitution elasticities across product types.

Using the CES structure of Feenstra (1994) and the estimation approach of Broda and Weinstein (2006), we combine benchmark U.S. product-level elasticities with newly constructed Japanese product-level estimates to provide cross-country evidence.

While both classifications exhibit ranking patterns broadly consistent with the theoretical ordering, only the IHY classification delivers statistically robust, distribution-wide, and economically meaningful support across all elasticity measures.

In contrast, support for the Rauch classification is fragile and sensitive to specification. These findings underscore the importance of transparent, HS-compatible, and structurally validated classification systems in empirical international trade research.

*Corresponding author. Nihon University, College of Economics, Tokyo, Japan; Research Institute for Impediments to Trade (RIIT). Email: naohiko.ijiri@riit.or.jp

†Nihon University Graduate School of Economics

Keywords: Trade product-type classification; substitution elasticity; ordered alternative; HS 6-digit; reproducibility.

1 Introduction

Product-type classifications play a central role in empirical international trade research, yet they are typically treated as convenient exogenous covariates rather than as structural hypotheses that can be confronted with data. Because classification systems embed implicit assumptions about market organisation and substitutability, their empirical validity is critical for structural estimation, quantitative welfare analysis, and policy evaluation.

Among existing frameworks, the classification proposed by Rauch (1999) has been particularly influential. Constructed using SITC codes, it categorises products as organised-exchange, reference-priced, or differentiated. Despite its widespread adoption, the Rauch classification raises concerns regarding transparency, reproducibility, and compatibility with modern HS-based trade data.

A further concern is conceptual in nature: although the classification is implicitly motivated by differences in substitutability across product types, it is rarely evaluated against independently estimated substitution elasticities. Most studies instead take the categories as given and examine downstream outcomes such as trade duration, margins of adjustment, or unit-value dispersion. As a result, the structural coherence between classification and the elasticity parameter that underpins quantitative trade models remains largely unexamined.

These issues are particularly salient in contemporary empirical applications. First, because the Rauch classification is SITC-based while modern detailed trade data are organised under the Harmonized System (HS), empirical applications require concordance procedures that can introduce noise and reduce reproducibility. Second, the assignment of reference-priced products relies partly on qualitative information that is not fully codified. More fundamentally, it remains unclear whether the implied ordering of substitution elasticities is borne out in the data.

To address these concerns, Ijiri et al. (2013) and Ijiri and Haneda (2022) propose a rule-based HS 6-digit classification (hereafter, the IHY classification). The IHY framework categorises products into Organised-exchange (O), Standardised differentiated (S), and Non-standardised differentiated (D) products based on observable institutional criteria such as exchange listings and internationally recognised standards. Unlike the Rauch typology, the IHY classification is explicitly designed to ensure HS compatibility, transparency, and reproducibility.

The central contribution of this study is to evaluate competing product-type classifications using substitution elasticity as a structural identification criterion. Because substitution elasticity

governs the curvature of CES demand and thus the intensity of competitive interaction across varieties, it provides a direct, theory-consistent benchmark for assessing whether classification systems capture economically meaningful differences in market organisation. Under a CES demand system, substitution elasticity measures the degree of replaceability across country varieties within a product. If a classification captures systematic differences in market organisation and standardisation, it should induce a clear ordered ranking of substitution elasticities across product types.

Formally, the theoretical hypotheses can be expressed as ordered alternatives:

$$\text{IHY} : O > S > D, \qquad \text{Rauch} : w > r > n.$$

That is, organised-exchange products are expected to exhibit the highest substitution elasticities, followed by intermediate categories, with the most differentiated products exhibiting the lowest elasticities.

Our empirical evaluation relies on two independent sources of elasticity estimates. First, we use the U.S. product-level substitution elasticities reported by Broda and Weinstein (2006), denoted σ_g^{US}.[1] Second, we construct new product-level substitution elasticity estimates for Japan using HS 6-digit import data within the same Feenstra–Broda–Weinstein framework. To assess robustness, we implement three alternative constructions—σ_g^{CBR}, σ_g^{FR}, and σ_g^{PSR}—that differ in their reference-country assumptions. Details are provided in Ijiri and Chen (2026) and summarised in Appendix C.

The empirical analysis proceeds in three stages. First, descriptive statistics assess whether theoretical ranking patterns appear in means and medians. Second, rank-based nonparametric tests designed for ordered alternatives assess monotone shifts across the full distribution. Third, regression analysis and the distribution-free effect-size measure Cliff's δ quantify both statistical significance and economic magnitude.

The results reveal a clear contrast. While both classifications sometimes display average ranking patterns consistent with theoretical predictions, only the IHY classification delivers statistically robust, distribution-wide, and economically meaningful support for the predicted ordering across both the U.S. benchmark and the Japanese estimates. Support for the Rauch classification is comparatively weaker and less consistent across elasticity measures.

This paper makes three main contributions to the empirical international trade literature. First, it treats product-type classification not as an exogenous premise but as a testable struc-

[1]The U.S. elasticity estimates are obtained from the publicly available replication files of Broda and Weinstein (2006), available at `https://www.columbia.edu/~dew35/TradeElasticities/TradeElasticities.html`. Accessed December 10, 2025.

tural hypothesis. Second, by combining the U.S. benchmark elasticities with independently constructed Japanese elasticities, it provides cross-country validation. Third, by integrating ordered-alternative testing with nonparametric magnitude measures, it offers a general empirical framework for evaluating theoretically motivated ranking structures. In doing so, it complements existing elasticity estimation work by turning these estimates toward the evaluation of classification systems themselves.

The remainder of the paper proceeds as follows. Section 2 reviews the conceptual foundations of the Rauch and IHY classifications and clarifies the structural channel linking market organisation to substitution elasticities. Section 3 outlines the estimation framework and introduces the elasticity measures. Section 4 presents the comparative empirical results. Section 5 concludes.

2 Related Literature and Structural Foundations of Product-Type Classification

This section situates the present study within the empirical international trade literature. It then clarifies the structural mechanism linking market organisation to substitution elasticity under a CES demand system, thereby motivating the ordered hypotheses tested in subsequent sections.

2.1 Product-Type Classification in Empirical Trade Research

The modern literature on product-type classification originates with Rauch (1999) , who emphasises the role of search costs and informational frictions in international trade.[2] Using the Standard International Trade Classification (SITC), Rauch categorises traded goods into organised-exchange, reference-priced, and differentiated products. The underlying premise is that differences in price observability and market organisation reflect systematic differences in search frictions and, ultimately, in effective substitutability.

Since its introduction, the Rauch classification has become a common empirical benchmark for capturing product heterogeneity. For example, Besedeš and Prusa (2006) show that the duration of trade relationships differs across product types; Rauch and Watson (2003) analyse matching and network formation in differentiated product trade; and Bastos and Silva (2010) relate product classification to vertical differentiation and unit-value dispersion. In much of this literature, product-type classification enters as an exogenous partition used to explain

[2]The Rauch product classification is publicly available from James Rauch's official website: `https://econweb.ucsd.edu/~jrauch/rauch_classification.html`. This paper uses the *Liberal* classification, which assigns products to organised-exchange or reference-priced categories whenever supporting evidence exists, whereas the *Conservative* classification assigns products to these categories only when evidence is unequivocal and otherwise classifies them as differentiated. Accessed December 10, 2025.

downstream outcomes.

Despite this widespread use, relatively little work evaluates whether product-type classifications correspond to independently estimated substitution elasticities. Existing studies typically condition on the classification and examine trade outcomes. Yet the structural coherence between classification and the substitution elasticity parameter that underpins modern quantitative trade models remains largely untested.[3]

In parallel, a substantial literature has developed methods for estimating product-level substitution elasticities within CES demand systems (Feenstra, 1994; Broda and Weinstein, 2006). However, these estimates are rarely used to validate classification systems themselves. This paper bridges these strands by treating product-type classification as a structural hypothesis about ordered substitution elasticities.

2.2 Market Organisation, CES Curvature, and Structural Implications

Under a CES demand system, the elasticity of substitution governs the curvature of the subutility aggregator and thus the intensity of competition across varieties. Within a narrowly defined product, σ_g determines how strongly relative demand responds to relative price changes across exporting countries.

Market organisation can affect effective substitution elasticities by shaping information sets, comparability of characteristics, and the dispersion of quality-adjusted prices. Public price quotation reduces informational asymmetries. International standards harmonise specifications, and formal exchanges provide common trading platforms and contract enforcement. These institutional features compress effective horizontal differentiation and increase perceived interchangeability across suppliers.

Conversely, when prices are not publicly observable, standards are weak or absent, and product characteristics are heterogeneous, buyers face greater search frictions and relationship-specific investments. Effective substitutability across suppliers is then lower.

[3] A separate line of research identifies product differentiation from observable product attributes or price information rather than institutional criteria. Bernini et al. (2018) propose the "Micro-D" classification, which uses the packaging (container) size observed at the finest level of customs nomenclature—Argentina's 12-digit SIM classification—as a proxy for differentiation, treating goods exported in small packages as differentiated and those shipped in bulk or large volumes as homogeneous; this is particularly effective at identifying differentiated products in the food and beverage sectors, which the SITC-based Rauch classification tends to miss. The Micro-D scheme, however, relies on a single country's detailed nomenclature and cannot be applied uniformly at the HS 6-digit level across countries, and its validity is corroborated only indirectly, through the relationship between unit-value levels and destination-market income. Bernard et al. (2015), by contrast, use the dispersion of export unit values within a given HS 6-digit product (a price-dispersion coefficient) as a proxy, treating products with smaller dispersion as more homogeneous; this is useful for measuring the degree of differentiation continuously within the relatively coarse HS 6-digit level. Both approaches, however, aim to identify or measure the degree of differentiation using observable proxies, rather than to test whether the ordered ranking of substitution elasticities implied by a classification holds statistically. For this reason, they are not the focus of the present analysis.

It follows that product-type classifications implicitly partition products according to systematic differences in the curvature of demand. If classification categories capture economically meaningful variation in market organisation, they should induce an ordered ranking in substitution elasticities. This implication can be expressed as the ordered restriction

$$\sigma^{(1)} > \sigma^{(2)} > \sigma^{(3)}.$$

The empirical analysis evaluates whether competing classification frameworks satisfy this restriction.

2.3 The Rauch Classification

The classification proposed by Rauch (1999) distinguishes products according to the informational environment in which trade occurs. Homogeneous products are subdivided into organised-exchange and reference-priced products.

Organised-exchange products are traded on formal commodity exchanges with publicly quoted prices and clearly defined standards. High price transparency and standardisation imply relatively high substitution elasticities.

Reference-priced products are not traded on formal exchanges but have benchmark prices reported in trade publications or industry sources. Although price signals are publicly available, standardisation may be weaker than in organised exchanges, implying intermediate substitution elasticities.

Differentiated products lack publicly observable reference prices and exhibit substantial heterogeneity across suppliers. Search frictions and relationship-specific investments are more pronounced, implying lower substitution elasticities.

The Rauch framework therefore implies the ordering

$$w > r > n.$$

However, it is constructed at the SITC level and partly relies on qualitative sources. Mapping to HS-based datasets requires concordance procedures, and the criteria used to identify reference-priced products are not fully codified.

2.4 The IHY Classification

The IHY classification (Ijiri et al., 2013; Ijiri and Haneda, 2022) adopts a rule-based and institutionally grounded approach constructed directly at the HS 6-digit level.

Under the IHY classification, homogeneous products are defined exclusively as organised-exchange products. These correspond to products traded in internationally regulated commodity markets under formal exchange mechanisms involving physical delivery. In such markets, product specifications are highly standardised and transaction prices are formed in organised exchanges, resulting in a high degree of price observability.

While these products share institutional characteristics with organised-exchange products in the Rauch classification, the IHY classification adopts a stricter criterion by limiting homogeneous products to transactions involving physical delivery. Under such institutional conditions, the transparency of prices and the standardisation of product specifications reduce the scope for effective supplier differentiation, leading to relatively high substitution elasticity σ_g. In this sense, organised-exchange products in the IHY classification are structurally consistent with the concept of homogeneous goods in the Rauch classification.

Differentiated products are subdivided into two groups.

Standardised differentiated products are subject to internationally recognised technical standards (e.g. ISO), which harmonise product specifications across suppliers without necessarily involving organised exchange trading. These standards enhance comparability across suppliers and improve the observability of product characteristics, while still allowing for some degree of supplier-specific differentiation. As a result, substitution across suppliers is feasible but imperfect, and substitution elasticities are expected to be intermediate between homogeneous and non-standardised differentiated products.

Non-standardised differentiated products lack internationally harmonised technical standards and exhibit substantial heterogeneity in product characteristics. In these markets, price information and product attributes are less easily comparable across suppliers, and differentiation plays a more prominent role in shaping demand. Consequently, substitution across suppliers is more limited, and substitution elasticities are expected to be relatively low.

The IHY framework implies the ordered structure

$$O > S > D.$$

Unlike the Rauch framework, the IHY classification is explicitly codifiable, HS-compatible, and reproducible across datasets, making it particularly suited to validation using HS 6-digit product-level elasticities.

Table 1 summarises the structural foundations of the Rauch and IHY classifications and highlights the channels through which each framework implies ordered differences in substitution elasticities.

Table 1: Structural Comparison of the Rauch and IHY Product-Type Classifications

	Rauch Classification	**IHY Classification**
Foundational Criterion	Informational environment and price reporting practices	Observable institutional characteristics (exchange listing, international standards)
Level of Construction	SITC-based aggregation	HS 6-digit-based rule-based classification
Homogeneous Segment	Organised-exchange and reference-priced goods	Organised-exchange goods traded in internationally regulated markets
Differentiated Segment	Differentiated products (residual category)	Standardised differentiated and non-standardised differentiated products
Primary Structural Channel	Price observability and information transparency	Degree of standardisation and institutional codification
Mechanism Affecting σ	Public price signals reduce perceived heterogeneity and search costs	Standardisation compresses horizontal differentiation across suppliers
Reproducibility	Partly qualitative classification criteria	Fully codifiable and rule-based assignment
HS Compatibility	Requires SITC–HS concordance	Directly constructed at the HS 6-digit level
Testable Ordered Hypothesis	$w > r > n$	$O > S > D$

Notes: The table contrasts the structural foundations of the two classification systems. Both frameworks imply ordered restrictions on product-level substitution elasticities σ_g. The empirical analysis evaluates whether independently estimated elasticities satisfy these ordered hypotheses.

The key distinction is that the Rauch framework infers product "homogeneity" from observed price-reporting practices, whereas the IHY classification explicitly codifies institutional standardisation at the HS 6-digit level. If substitution elasticities capture the curvature of the CES aggregator, classifications aligned with the underlying sources of effective competition should generate stable ordered restrictions across countries and across elasticity constructions. The next section therefore turns to the identification of product-level substitution elasticities and formulates these ordered restrictions as econometric hypotheses.

2.5 Classification Concordance and Data

To compare the Rauch and IHY classifications within a common HS 6-digit-level dataset, we map the SITC-based Rauch categories to HS 6-digit products using official SITC–HS concordance tables[4]. When multiple SITC categories correspond to a single HS 6-digit product, we assign the dominant Rauch category based on the modal classification across concordance mappings, following standard practice in the empirical trade literature.

The empirical analysis uses Japanese import data at the HS 6-digit level, including annual

[4]The official SITC–HS concordance tables are provided by the United Nations Statistics Division (UNSD). See: `https://unstats.un.org/unsd/classifications/Econ`. Accessed December 10, 2025.

import values and quantities by product and trading partner. Substitution elasticities estimated from these data are combined with the U.S. elasticity estimates from Broda and Weinstein (2006), allowing us to evaluate classification validity independently of country-specific measurement issues.

3 Estimation of Substitution Elasticities and Empirical Strategy

This section describes the structural framework used to estimate substitution elasticities and explains how these estimates are used to evaluate competing product-type classification systems. We first clarify why substitution elasticity provides a structural criterion for product differentiation, then outline the Feenstra–Broda–Weinstein framework, describe the construction of Japanese elasticity measures, and finally present the empirical strategy for classification validation.

3.1 Substitution Elasticity as a Structural Criterion

For a product-type classification to carry economic content, it should correspond to systematic differences in substitutability across categories. Both the Rauch and IHY frameworks are motivated by this premise: differences in market organisation, price transparency, and standardisation are expected to map into differences in substitution elasticities.

Substitution elasticity provides a structural measure of effective product differentiation. Higher elasticities indicate that country varieties are close substitutes, consistent with homogeneous characteristics and transparent pricing. Lower elasticities indicate stronger differentiation and weaker substitutability.

This yields a clear and testable empirical implication: product types associated with more standardised and transparent market environments should exhibit higher substitution elasticities. Substitution elasticity therefore offers an objective criterion against which alternative classification systems can be evaluated.

3.2 The Feenstra (1994) and Broda–Weinstein (2006) Framework

Following Feenstra (1994) and Broda and Weinstein (2006), preferences over imported varieties within product g follow a constant elasticity of substitution (CES) structure. The subutility function is

$$U_g = \left(\sum_c q_{gc}^{\frac{\sigma_g - 1}{\sigma_g}} \right)^{\frac{\sigma_g}{\sigma_g - 1}}, \tag{3.1}$$

where q_{gc} denotes the quantity of product g imported from country c, and σ_g is the elasticity of substitution across varieties.

Under this demand system, we identify σ_g from joint variation in import prices and expenditure shares across supplying countries. Intuitively, products whose expenditure shares respond more strongly to relative price movements exhibit higher substitution elasticities. Empirically, elasticities are estimated product-by-product at the HS 6-digit level using country-level variation in import unit values and expenditure shares. Broda and Weinstein (2006) implement this approach using detailed U.S. trade data, producing benchmark estimates that are widely used in the literature.

Full derivations and identification conditions are provided in Appendix B.

3.3 Estimation of Substitution Elasticities for Japan

To complement the U.S. benchmark estimates, we use substitution elasticities constructed from Japanese import data within the same structural framework.

Ijiri and Chen (2026) estimate HS 6-digit product-level substitution elasticities for Japan. Building on their approach, we consider three alternative constructions that differ in their choice of reference country, which affects the measurement of relative price movements.

Under the **Product-Specific Reference (PSR)** specification, the country with the largest import share in the base year t_0 is selected as the reference for each product. This choice anchors relative prices to the dominant supplier and tends to maximise within-product variation.

Under the **Fixed Reference (FR)** specification, the United States is imposed as a common reference country across all products. This provides a uniform external benchmark and eliminates variation arising from endogenous reference-country choice.

Finally, the **Conservative Baseline Reference (CBR)** elasticity is defined as

$$\sigma_g^{CBR} = \min\left(\sigma_g^{PSR}\sigma_g^{FR}\right).$$

This CBR limits the influence of potentially upward-biased estimates associated with reference-country selection and serves as a robustness benchmark.

Using multiple constructions allows us to assess whether the performance of a classification system is robust to alternative elasticity measures. Formal definitions of σ_g^{PSR}, σ_g^{FR}, and σ_g^{CBR} are provided in Appendix C.

3.4 Empirical Strategy for Evaluating Product-Type Classification

The empirical strategy evaluates whether substitution elasticities follow the theoretically predicted ordering implied by each classification framework.

Under the IHY classification, the predicted ordering is

$$\begin{aligned}\text{Organised-exchange } (O) &> \text{Standardised differentiated } (S)\\ &> \text{Non-standardised differentiated } (D).\end{aligned}$$

Under the Rauch classification, the predicted ordering is

$$\begin{aligned}\text{Organised-exchange } (w) &> \text{Reference-priced } (r)\\ &> \text{Differentiated } (n).\end{aligned}$$

We evaluate these ordered hypotheses using three complementary approaches. First, we compare mean and median elasticities across product types. Second, we implement rank-based tests designed for ordered alternatives to assess monotone shifts across the distribution. Third, we use regression analysis and distribution-free effect-size measures to quantify magnitude and statistical significance. By combining elasticity estimates derived from independent data sources and alternative constructions, this framework provides a structurally grounded and empirically robust assessment of classification validity.

Importantly, the classification assignments are constructed independently of the elasticity estimates. This independence ensures that the empirical tests do not mechanically impose the predicted ordering and allows substitution elasticity to serve as an external structural validation criterion.

The next section presents the empirical results and compares the performance of the Rauch and IHY classification systems.

4 Empirical Results and Structural Comparison

This section evaluates the structural validity of the IHY and Rauch product-type classifications by examining whether substitution elasticities exhibit the theoretically predicted monotone ordering across product types. Because substitution elasticity is a structural parameter governing competitive interaction and demand substitutability, systematic ordering provides a direct test of whether a classification captures economically meaningful differences in market organisation.

The analysis uses four elasticity measures: the U.S. product-level elasticities reported by Broda and Weinstein (2006) (σ_g^{US}) and three measures estimated from Japanese import data (σ_g^{CBR}, σ_g^{FR}, and σ_g^{PSR}).

The theoretical predictions are:

$$\text{IHY classification:} \quad \sigma_O > \sigma_S > \sigma_D, \tag{4.1}$$

$$\text{Rauch classification:} \quad \sigma_w > \sigma_r > \sigma_n. \tag{4.2}$$

We evaluate these predictions using descriptive statistics, rank-based trend tests, permutation-based Jonckheere–Terpstra tests, parametric regressions, and nonparametric effect-size measures.

4.1 Descriptive Evidence

Panel A of Tables 2 and 3 reports mean and median substitution elasticities by product type.

Under the IHY classification, median elasticities satisfy the predicted ordering $\sigma_O > \sigma_S > \sigma_D$ for all four elasticity measures. Mean elasticities satisfy the ordering for σ_g^{US} and σ_g^{CBR}. Although the mean ranking deviates for σ_g^{FR} and σ_g^{PSR}, the median ranking remains fully consistent, indicating that the monotone structure is robust in the central part of the distribution.

In contrast, the Rauch classification exhibits weaker and less stable ranking patterns. Mean rankings frequently violate the predicted ordering, particularly for the Japanese elasticity measures, suggesting a looser correspondence between Rauch categories and underlying substitutability.

These distributional patterns are visualised in Figures 1 and 2.

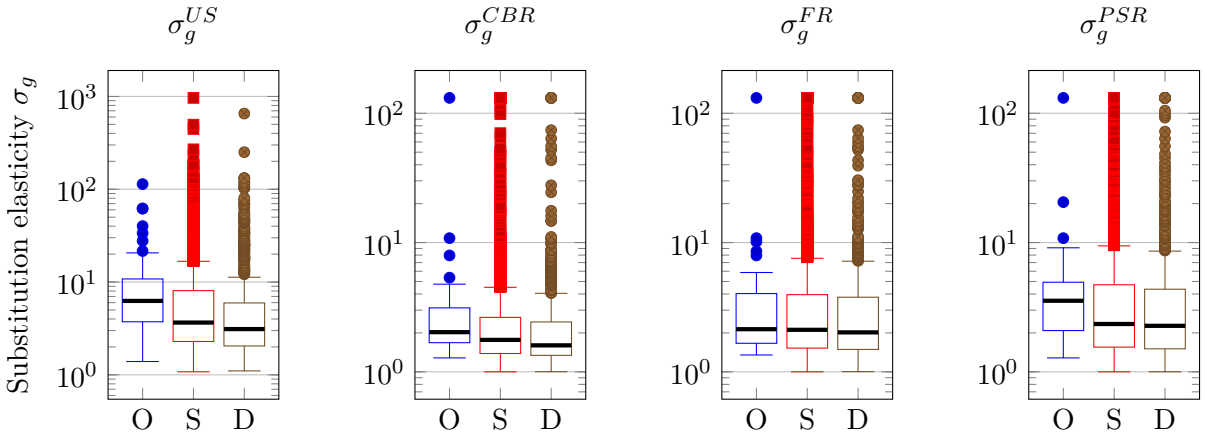

Figure 1: Distribution of substitution elasticities across product types under the IHY classification. O denotes organised-exchange products, S standardised differentiated products, and D non-standardised differentiated products. Boxes represent interquartile ranges, horizontal lines denote medians, and whiskers extend to adjacent observations. The vertical axis is shown on a logarithmic scale.

Figures 1 and 2 provide graphical representations of the distribution of substitution elasticities across product types under the IHY and Rauch classifications, respectively. Several important patterns emerge.

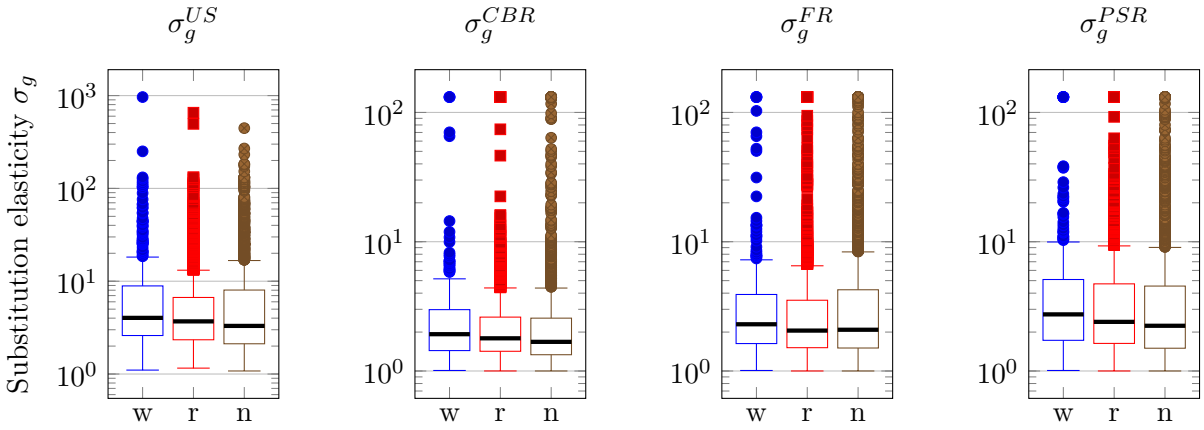

Figure 2: Distribution of substitution elasticities across product types under the Rauch classification. w denotes organised-exchange products, r reference-priced products, and n differentiated products. Boxes represent interquartile ranges, horizontal lines denote medians, and whiskers extend to adjacent observations. The vertical axis is shown on a logarithmic scale.

First, under the IHY classification (Figure 1), the distributions exhibit a clear and systematic ordering consistent with theoretical predictions. For all four elasticity measures, the median and interquartile ranges follow the ranking $O > S > D$. In addition, the distributions are well separated (IQR overlap <10% for US estimates), with limited overlap between adjacent categories, particularly between organised and differentiated products. This pattern indicates that the IHY classification aligns closely with structural differences in substitutability implied by the CES demand framework.

Second, under the Rauch classification (Figure 2), the distributions display greater overlap across categories. Although homogeneous products tend to exhibit higher elasticities on average, the separation between reference-priced and differentiated products is less pronounced. The interquartile ranges overlap extensively, and the ranking structure is less clearly defined.

These figures complement the formal statistical tests in Tables 2 and 3 by visually confirming that the IHY classification yields a clearer and more stable ordering of substitution elasticities than the Rauch classification, thereby reinforcing the conclusion that the IHY framework possesses stronger structural validity.

4.2 Rank Trend Tests: Cuzick Test

Panel B of Tables 2 and 3 reports Cuzick's rank trend test (Cuzick, 1985), which assesses monotone trends across ordered groups.

For the IHY classification, the null of identical distributions is rejected for all elasticity measures. The z-statistics are positive and statistically significant across σ_g^{US}, σ_g^{CBR}, σ_g^{FR}, and σ_g^{PSR}, consistent with increasing substitution elasticities from D to S to O.

By contrast, the Rauch classification yields less robust evidence. While σ_g^{US}, σ_g^{CBR}, and σ_g^{PSR} reject the null, the FR elasticity does not ($p = 0.843$), indicating sensitivity across elasticity constructions.

Table 2: Substitution Elasticity Ordering under the IHY Classification

	σ_g^{US}	σ_g^{CBR}	σ_g^{FR}	σ_g^{PSR}
Panel A. Descriptive Statistics				
Mean ranking	O>S>D	O>S>D	S>O>D	S>O>D
Median ranking	O>S>D	O>S>D	O>S>D	O>S>D
Observations	4,293	3,831	3,725	3,506
Panel B. Cuzick Rank Trend Test				
z-statistic	5.93	3.49	2.05	2.32
p-value	0.000***	0.000***	0.040**	0.020**

Notes: Cuzick (1985) rank trend test evaluates the ordered alternative $O > S > D$. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 3: Substitution Elasticity Ordering under the Rauch Classification

	σ_g^{US}	σ_g^{CBR}	σ_g^{FR}	σ_g^{PSR}
Panel A. Descriptive Statistics				
Mean ranking	w>n>r	n>w>r	n>w>r	n>w>r
Median ranking	w>r>n	w>r>n	w>r>n	w>r>n
Observations	4,267	3,969	3,701	3,491
Panel B. Cuzick Rank Trend Test				
z-statistic	3.94	4.85	0.20	3.77
p-value	0.000***	0.000***	0.843	0.000***

Notes: Cuzick (1985) rank trend test evaluates the ordered alternative $w > r > n$. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

4.3 Primary Structural Test: Jonckheere–Terpstra Permutation Test

Table 4 reports permutation-based Jonckheere–Terpstra (JT) tests ($B = 10{,}000$ permutations) (Jonckheere, 1954; Terpstra, 1952; Hollander et al., 2014). Because the JT test is designed for ordered alternatives and relies on the full distribution, it provides a direct nonparametric assessment of the structural ranking implied by each classification.

For the IHY classification, the null of identical distributions is rejected for all elasticity measures (σ_g^{US}: $p < 0.001$; σ_g^{CBR}: $p = 0.0001$; σ_g^{FR}: $p = 0.0243$; σ_g^{PSR}: $p = 0.0138$), providing strong evidence consistent with monotone ordering across independent elasticity measures and across the U.S. and Japanese datasets.

For the Rauch classification, ordered relationships are detected for several elasticity measures but are weaker and less consistent than under the IHY classification. The null hypothesis is rejected for σ_g^{US} ($p = 0.0002$) and σ_g^{CBR} ($p < 0.001$) and σ_g^{PSR} ($p < 0.001$). However, under the FR specification, the z-statistic is -0.27 ($p = 0.609$), indicating that the null is not rejected and

that the ordering is in fact reversed. This result suggests that, under the FR elasticity measure, the Rauch classification categories show no systematic separation in substitution elasticities—and if anything a reversal—which runs counter to the theoretical prediction.

Moreover, the magnitude and consistency of the JT statistics are substantially stronger under the IHY classification. These results indicate that while the Rauch classification captures some systematic variation in substitution elasticities under certain specifications, its support is sensitive to elasticity construction and fragile in the FR case. Table 4 confirms this quantitatively: JT and z statistics are uniformly larger and more consistently significant under the IHY classification, indicating stronger and more robust alignment between classification categories and substitution-elasticity structure.

The JT results thus confirm that the IHY ordering is stable and consistently supported across elasticity measures, whereas the Rauch ordering is weaker, less consistent, and sensitive to specification.

Table 4: Permutation-Based Jonckheere–Terpstra Tests

	IHY Classification			Rauch Classification		
Elasticity Measure	JT	Z	p-value	JT	z	p-value
σ_g^{US}	1,888,702	5.81	0.0000***	2,652,841	3.69	0.0002***
σ_g^{CBR}	1,317,370	3.60	0.0001***	2,323,340	4.78	0.0000***
σ_g^{FR}	1,332,405	2.02	0.0243**	1,848,531	−0.27	0.6089
σ_g^{PSR}	1,196,954	2.23	0.0138**	1,760,705	3.85	0.0000***

Notes: This table reports permutation-based Jonckheere–Terpstra tests of ordered alternatives under the IHY classification ($O > S > D$) and the Rauch classification ($w > r > n$). Permutation p-values are based on $B = 10{,}000$ random permutations with random seed fixed at 42 to ensure reproducibility. The JT statistic counts the number of pairwise comparisons consistent with the predicted ordering. The z-statistic is an empirical standardisation based on the permutation mean and standard deviation (not the large-sample asymptotic formula); the two are numerically close in all cases. A negative Z under Rauch–σ_g^{FR} ($Z = -0.27$, $p = 0.609$) indicates that the ordering is reversed relative to the theoretical prediction under this specification. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

4.4 Parametric Evidence: Regression Analysis

To quantify the magnitude of monotone relationships, we estimate

$$\ln \sigma_g = \alpha + \beta \cdot \text{Score}_g + \varepsilon_g, \tag{4.3}$$

where Score_g assigns increasing numerical values according to the theoretical ordering (IHY: $D = 1$, $S = 2$, $O = 3$; Rauch: $n = 1$, $r = 2$, $w = 3$).

Because the classification is defined independently of the elasticity estimates, the ordered score is predetermined with respect to σ_g, allowing β to capture whether independently assigned

classification categories align systematically with substitution elasticity differences, rather than implying a causal effect of classification itself.

Table 5 reports the results. Under the IHY classification, the ordered-score coefficient is positive across all four elasticity measures and statistically significant for three of them (σ_g^{US}, σ_g^{FR}, and σ_g^{PSR}). The σ_g^{CBR} estimate is also positive but is not statistically significant at conventional levels. For σ_g^{US}, $\hat{\beta} = 0.160$ implies that moving up by one classification category is associated with an increase in substitution elasticity of approximately $e^{0.160} - 1 \approx 17\%$, reflecting economically meaningful structural separation across classification categories. The results are broadly similar across elasticity measures, although the strength of statistical evidence differs across specifications.

Table 5: Regression Results: Ordered Product-Type Score and Substitution Elasticity

	σ_g^{US}	σ_g^{CBR}	σ_g^{FR}	σ_g^{PSR}
IHY Classification				
Score coefficient (β)	0.160***	0.033	0.089***	0.080**
Standard error	(0.030)	(0.022)	(0.033)	(0.037)
Observations	4,293	3,995	3,725	3,506
Rauch Classification				
Score coefficient (β)	0.010	0.009	-0.067**	0.020
Standard error	(0.021)	(0.018)	(0.031)	(0.027)
Observations	4,267	3,969	3,701	3,491

Notes: Dependent variable is $\ln \sigma_g$. Score is coded according to the theoretical ordering (IHY: D=1, S=2, O=3; Rauch: n=1, r=2, w=3).
Robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

In contrast, the Rauch classification yields coefficients close to zero. Notably, the FR specification produces a statistically significant negative coefficient, which runs counter to the theoretical ordering.

Figure 3 summarises these patterns.

Coefficients under the IHY classification are consistently positive, whereas Rauch coefficients are smaller and include both small positive estimates and a negative estimate. The regression evidence thus provides broad parametric support that the IHY score is associated with higher substitution elasticities, although the strength of statistical evidence varies across elasticity constructions. By contrast, the Rauch score does not show a robust positive association. These results suggest that the IHY classification better captures systematic variation in demand curvature than the Rauch classification.

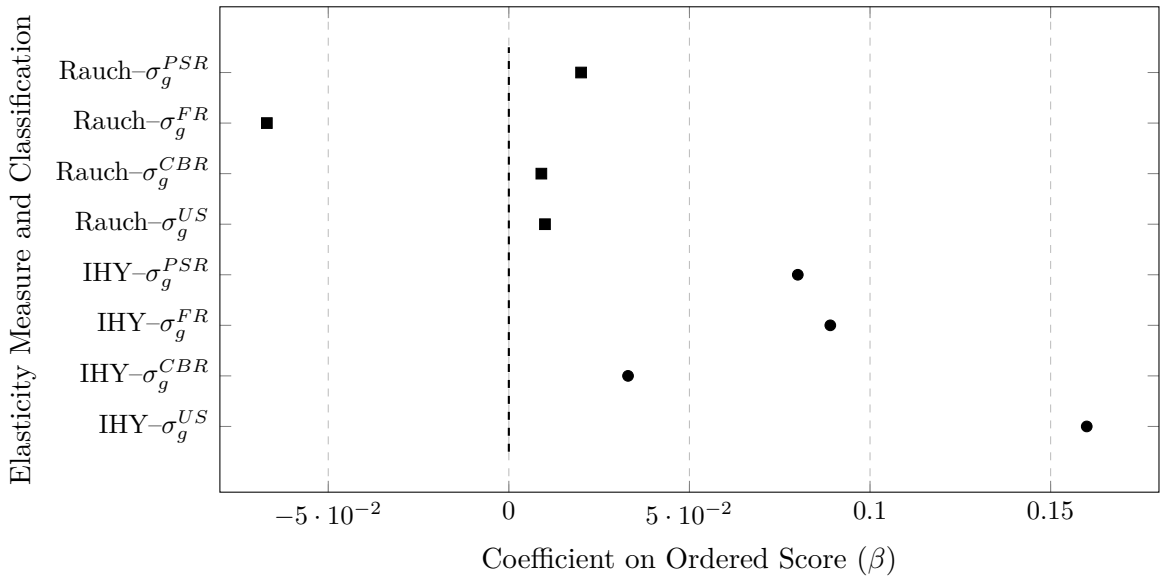

Figure 3: Regression coefficients of ordered product-type score on $\ln \sigma_g$.

Notes: The figure plots the OLS coefficient $\hat{\beta}$ from equation (4.3) for each elasticity measure. Circles denote estimates under the IHY score ($D = 1$, $S = 2$, $O = 3$) and squares denote estimates under the Rauch score ($n = 1$, $r = 2$, $w = 3$). The dashed vertical line indicates zero. Elasticity measures include U.S. estimates (Sigma) and Japanese estimates (CBR, FR, and PSR).

4.5 Effect Size and Economic Magnitude

Table 6 reports Cliff's δ (Cliff, 1993; Vargha and Delaney, 2000), a distribution-free effect-size measure.

Effect sizes under the IHY classification are uniformly positive and economically meaningful. For example, the organised-exchange versus differentiated comparison yields $\delta = 0.407$ for σ_g^{US}, indicating substantial separation in the elasticity distribution.

Effect sizes under the Rauch classification are smaller and less consistent, including a negative comparison under the FR elasticity. These patterns indicate that substitutability differences are economically substantial under the IHY classification but weaker under the Rauch classification. The effect-size evidence complements the JT tests and regression results by quantifying the economic magnitude of the implied structural differences.

4.6 Summary and Structural Interpretation

Across descriptive statistics, rank-based tests, permutation-based structural tests, regression analysis, and effect-size measures, the evidence consistently supports the structural validity of the IHY classification more strongly than that of the Rauch classification.

Three findings stand out. First, substitution elasticities under the IHY classification exhibit stable monotone ordering across independent elasticity measures and across the U.S. and Japanese

Table 6: Effect Size Analysis Across Product Types (Cliff's δ)

	IHY Product-Type Classification (O–S–D)		Rauch Product-Type Classification (w–r–n)	
	δ	Comparison	δ	Comparison
σ_g^{US}				
O vs D / w vs n	0.407	O–D	0.120	w–n
O vs S / w vs r	0.293	O–S	0.088	w–r
S vs D / r vs n	0.106	S–D	0.041	r–n
σ_g^{CBR}				
O vs D / w vs n	0.303	O–D	0.144	w–n
O vs S / w vs r	0.240	O–S	0.094	w–r
S vs D / r vs n	0.059	S–D	0.061	r–n
σ_g^{FR}				
O vs D / w vs n	0.108	O–D	0.053	w–n
O vs S / w vs r	0.061	O–S	0.094	w–r
S vs D / r vs n	0.041	S–D	−0.034	r–n
σ_g^{PSR}				
O vs D / w vs n	0.270	O–D	0.104	w–n
O vs S / w vs r	0.228	O–S	0.048	w–r
S vs D / r vs n	0.037	S–D	0.063	r–n

Notes: Cliff's δ is defined as $P(X > Y) - P(X < Y)$, where X and Y denote observations from the higher- and lower-ranked categories, respectively.
Positive values support the predicted ordering (IHY: O>S>D; Rauch: w>r>n).
Effect sizes are computed from the Mann–Whitney statistic (Mann and Whitney, 1947) using midrank adjustments.

datasets; the permutation-based JT tests provide particularly direct support. Second, evidence for the Rauch ordering is sensitive to the elasticity construction: under the FR measure the z-statistic is negative ($Z = -0.27$, $p = 0.609$), indicating a reversal of the predicted ordering rather than mere non-significance, and mean rankings frequently violate the predicted ordering for the Japanese estimates. Third, regression and effect-size analyses indicate that the IHY ordering is associated with economically meaningful differences in substitutability, whereas the corresponding differences under the Rauch classification are smaller and less stable.

These results consistently support the conclusion that the HS-based IHY classification more reliably captures structural heterogeneity in product substitutability and therefore provides a more robust foundation for empirical trade analysis. These findings motivate the implications discussed in Section 5.

5 Conclusion

This study reassesses the structural validity of two influential product-type classification systems—the Rauch (1999) typology and the HS 6-digit-based IHY framework—by evaluating their consistency with independently estimated substitution elasticities. Rather than treating product-type classifications as exogenously given descriptors, the paper examines whether they align with a key structural parameter governing international trade: the elasticity of substitution.

Using the U.S. elasticities reported by Broda and Weinstein (2006) together with three Japanese elasticity measures (σ_g^{CBR}, σ_g^{FR}, and σ_g^{PSR}), we implement a validation strategy combining (i) descriptive ranking comparisons, (ii) distribution-wide ordered-alternative tests, and (iii) regression and nonparametric effect-size analyses. This multi-layered approach allows us to assess not only central tendencies but also the robustness and economic magnitude of ordering across the full elasticity distribution and across countries.

The evidence provides strong and consistent support for the IHY classification. Across all four elasticity measures, the IHY classification aligns with the predicted ordering

Organised-exchange (O) > Standardised differentiated (S) > Non-standardised differentiated (D).

This ordering holds at the median and is supported by distribution-wide ordered-alternative tests. Moreover, differences across product types are economically meaningful, with large and stable effect sizes for comparisons between organised-exchange and differentiated products. These results indicate that the IHY classification captures systematic differences in substitutability consistent with its institutional foundations.

In contrast, support for the Rauch classification is weaker and less consistent. Although several elasticity measures support the predicted ordering, the evidence is fragile: under the FR specification the predicted ordering is reversed ($Z = -0.27$, $p = 0.609$), and effect sizes are generally smaller than under the IHY classification. This pattern suggests that the Rauch categories align less consistently with substitution-elasticity structure.

These findings have important methodological implications. First, product-type classifications embody implicit structural assumptions about substitutability and market organisation, and their validity should therefore be evaluated empirically rather than assumed. Second, rigorous validation requires both distribution-wide tests and effect-size assessments in addition to mean or median comparisons. The framework developed here—combining ordered-alternative tests with nonparametric effect-size evaluation—provides a general approach for assessing theoretically motivated ranking structures in empirical classification systems.

The results also have substantive implications for empirical trade research. Because substitution elasticities play a central role in quantitative trade models and welfare analysis, classification systems that align more closely with independently estimated elasticity structure provide a more reliable empirical foundation for both positive and policy-oriented work. The superior performance of the IHY classification suggests that HS-based classifications grounded in transparent institutional criteria more accurately reflect structural heterogeneity in international trade.

Several limitations warrant further investigation. Elasticity estimates are subject to measurement error and may vary across sample periods and estimation methods. Future work incorporating additional countries, time periods, and alternative identification strategies would further strengthen external validity.

More broadly, this study contributes to the empirical trade literature by demonstrating that product-type classification systems can be evaluated using independently estimated structural parameters such as substitution elasticities. This approach provides a general empirical framework for assessing whether classification schemes align with theoretically meaningful sources of economic heterogeneity and demand structure.

Ultimately, these findings indicate that product-type classification is not merely a descriptive convention but a structural component of empirical trade analysis. When classification systems differ in their alignment with substitution elasticities, they imply different interpretations of product heterogeneity and market organisation; the evidence presented here favours the HS-based IHY framework over the traditional Rauch typology.

A Reproducibility of the Rauch Product Classification

This appendix documents the data sources and procedures used to implement the Rauch (1999) product classification in order to ensure full reproducibility.

The original Rauch classification is publicly available from James Rauch's official website:

`https://econweb.ucsd.edu/~jrauch/rauch_classification.html`

The dataset provides two alternative versions of the classification: the *Conservative* and the *Liberal* classifications. The Conservative classification assigns products to organised-exchange or reference-priced categories only when there is unequivocal evidence of exchange trading or publicly quoted reference prices. In contrast, the Liberal classification assigns products to these categories whenever supporting evidence exists, resulting in a broader assignment of products to organised-exchange and reference-priced categories and a smaller residual differentiated category.

Following the standard practice in the empirical trade literature, this paper adopts the Liberal classification. This choice reflects the objective of capturing systematic differences in market organisation and price transparency while avoiding excessive classification into the residual differentiated category due to incomplete documentation.

Because the original Rauch classification is defined at the SITC level, we convert the classification to HS 6-digit products using official SITC–HS concordance tables provided by the United Nations Statistics Division (UNSD):

`https://unstats.un.org/unsd/classifications/Econ`

When multiple SITC categories correspond to a single HS 6-digit product, we assign the dominant Rauch category based on the modal mapping across concordance entries. This procedure ensures consistency between classification assignment and the HS-based elasticity estimates used in the empirical analysis.

All classification files, concordance tables, and processed HS-level classification data used in this paper are archived by the authors and are available upon request. Because the original Rauch classification is publicly accessible and the concordance tables are official and stable, the classification procedure described above can be fully replicated.

This procedure follows standard practice in empirical trade studies implementing the Rauch classification using HS-based datasets (e.g., Besedeš and Prusa, 2006; Nunn, 2007).

B Theoretical Framework for Estimating Substitution Elasticities

This appendix summarises the theoretical identification of product-level substitution elasticities σ_g based on Feenstra (1994) and Broda and Weinstein (2006).

B.1 Demand System

For each HS 6-digit product g, preferences follow CES:

$$U_g = \left(\sum_{c \in C_g} q_{gc}^{\frac{\sigma_g - 1}{\sigma_g}} \right)^{\frac{\sigma_g}{\sigma_g - 1}} \tag{B.1}$$

Expenditure minimization yields:

$$q_{gc} = \left(\frac{p_{gc}}{P_g} \right)^{-\sigma_g} \frac{E_g}{P_g} \tag{B.2}$$

CES price index:

$$P_g = \left(\sum_c p_{gc}^{1-\sigma_g} \right)^{\frac{1}{1-\sigma_g}} \tag{B.3}$$

B.2 Share Equation

$$s_{gc} = \frac{p_{gc}^{1-\sigma_g}}{\sum_k p_{gk}^{1-\sigma_g}} \tag{B.4}$$

Log-difference:

$$\Delta \ln s_{gc} = (1 - \sigma_g) \Delta \ln p_{gc} - \Delta \ln P_g \tag{B.5}$$

Substitute price index:

$$\Delta \ln s_{gc} = (1 - \sigma_g) \left(\Delta \ln p_{gc} - \sum_k s_{gk} \Delta \ln p_{gk} \right) + \varepsilon_{gc} \tag{B.6}$$

B.3 Identification

$$\Delta \ln \tilde{p}_{gc} = \Delta \ln p_{gc} - \sum_k s_{gk} \, \Delta \ln p_{gk} \tag{B.7}$$

$$\Delta \ln s_{gc} = (1 - \sigma_g) \Delta \ln \tilde{p}_{gc} + \varepsilon_{gc} \tag{B.8}$$

$$\sigma_g = 1 - \frac{\mathrm{Cov}(\Delta \ln s_{gc}, \Delta \ln \tilde{p}_{gc})}{\mathrm{Var}(\Delta \ln \tilde{p}_{gc})} \tag{B.9}$$

This expression shows that substitution elasticity is identified from the covariance between expenditure-share changes and relative price changes across supplying countries, conditional on the CES demand structure. This identification strategy relies on cross-country variation in relative prices and assumes that supply-driven price movements are not perfectly correlated with demand shocks affecting expenditure shares. Following Feenstra (1994) and Broda and Weinstein (2006), instrumental-variable procedures and cross-sectional variation in supplier-specific price movements help isolate substitution responses.

C Estimation of Japanese Product-Level Elasticities

C.1 Product-Specific Reference (PSR)

$$c_g^* = \arg\max_c M_{gc,t_0} \tag{C.1}$$

$$\Delta \ln s_{gc,t} = (1-\sigma_g)\Delta \ln \left(\frac{p_{gc,t}}{p_{gc_g^*,t}} \right) + \nu_{gc,t} \tag{C.2}$$

C.2 Fixed Reference (FR)

$$\Delta \ln s_{gc,t} = (1-\sigma_g)\Delta \ln \left(\frac{p_{gc,t}}{p_{gUS,t}} \right) + \nu_{gc,t} \tag{C.3}$$

C.3 Conservative Baseline Reference (CBR)

$$\sigma_g^{CBR} = \min\left\{ \sigma_g^{PSR}, \sigma_g^{FR} \right\} \tag{C.4}$$

D Statistical Foundations

D.1 Jonckheere–Terpstra Test

The Jonckheere–Terpstra test (Jonckheere, 1954; Terpstra, 1952; Hollander et al., 2014) is a nonparametric test designed to detect ordered differences across multiple groups.

$$H_0 : \theta_1 = \theta_2 = \theta_3 \tag{D.1}$$

$$JT = \sum_{i<j} \sum_{x \in G_i} \sum_{y \in G_j} I(x > y) \tag{D.2}$$

D.2 Cliff's δ

$$\delta = P(X > Y) - P(X < Y) \tag{D.3}$$

$$\hat{\delta} = \frac{2U}{n_X n_Y} - 1 \tag{D.4}$$

E Empirical Implementation

$$\Delta \ln s_{gc,t} = \beta_g \Delta \ln \tilde{p}_{gc,t} + \varepsilon_{gc,t} \tag{E.1}$$

$$\sigma_g = 1 - \beta_g \tag{E.2}$$

Products with fewer than three exporters are excluded.

References

Bastos, P. and Silva, J. (2010). The quality of a firm's exports: Where you export to matters. *Journal of International Economics*, 82(2):99–111.

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

Bernini, F., González, J., Hallak, J. C., and Vicondoa, A. (2018). The micro-D classification: A new approach to identifying differentiated exports. *Economía*, 18(2):59–85.

Besedeš, T. and Prusa, T. J. (2006). Ins, outs, and the duration of trade. *Canadian Journal of Economics*, 39(1):266–295.

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

Cliff, N. (1993). Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin*, 114(3):494–509.

Cuzick, J. (1985). A Wilcoxon-type test for trend. *Statistics in Medicine*, 4(1):87–90.

Feenstra, R. C. (1994). New product varieties and the measurement of international prices. *American Economic Review*, 84(1):157–177.

Hollander, M., Wolfe, D. A., and Chicken, E. (2014). *Nonparametric Statistical Methods*. Wiley, 3rd edition.

Ijiri, N. and Chen, S. (2026). Estimating the Elasticity of Substitution for Traded Products: Evidence from Japan. RIIT Working Paper Series, RWP-2026-003.

Ijiri, N. and Haneda, S. (2022). A Note on the Concordance between the International Classification for Standards (ICS) and the Harmonised Commodity Description and Coding System (HS). RIIT Working Paper Series, RWP-2022-001.

Ijiri, N., Yamano, N., and Miao, G. (2013). International standards, product substitutability, and global trade network. *EURAS Proceedings*, pages 187–201.

Jonckheere, A. R. (1954). A distribution-free k-sample test against ordered alternatives. *Biometrika*, 41(1-2):133–145.

Mann, H. B. and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18(1):50–60.

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

Rauch, J. E. and Watson, J. (2003). Starting small in an unfamiliar environment. *International Journal of Industrial Organization*, 21(7):1021–1042.

Terpstra, T. J. (1952). The asymptotic normality and consistency of kendall's test against trend. *Nederl. Akad. Wetensch. Proc.*, 55:327–333.

Vargha, A. and Delaney, H. D. (2000). A critique and improvement of the CL common language effect size statistics of McGraw and Wong. *Journal of Educational and Behavioral Statistics*, 25(2):101–132.