# Price Homogeneity and Goods Classification: Technical Methodology

## Overview

This document describes a data-driven methodology for measuring price homogeneity in international trade and classifying goods into Rauch (1999) categories. The approach addresses four interconnected questions:

1. **Homogeneity**: How dispersed are prices across exporters within a product category?
2. **Divergence**: Do large and small exporters exhibit systematically different pricing?
3. **Classification**: Can we empirically identify homogeneous, reference-priced, and differentiated goods?
4. **Validation**: How do data-driven classifications compare to Rauch's (1999) original taxonomy?

---

## 1. Price Dispersion Measurement

### 1.1 Dominant Unit-of-Measure Restriction

Trade data often report quantities in multiple units of measure (UOM) within the same HS6 code (e.g., kilograms, items, litres). Comparing unit values across incompatible UOMs is meaningless, so we restrict analysis to the **dominant UOM**—the unit capturing the largest share of trade value within each HS6-year cell.

For each HS6 code $h$ and year $t$:

$$\text{UOM}^*_{ht} = \arg\max_u \sum_e V_{ehtu}$$

where $V_{ehtu}$ is trade value for exporter $e$ in UOM $u$. We compute **coverage** as:

$$\text{Coverage}_{ht} = \frac{\sum_e V_{eht,\text{UOM}^*}}{\sum_{e,u} V_{ehtu}}$$

Observations with coverage below a threshold (default: 85%) are flagged as potentially unreliable.

### 1.2 Exporter-Level Unit Values

Within the dominant UOM, we compute exporter-level unit values by aggregating across importers:

$$P_{eht} = \frac{\sum_i V_{eiht}}{\sum_i Q_{eiht}}$$

where $i$ indexes importers. This yields a single price observation per exporter-HS6-year.

### 1.3 Dispersion Metrics

We measure price dispersion using the **log interquartile ratio** between the 90th and 10th percentiles across exporters:

$$\text{LogGap}_{ht} = \log \left( \frac{P_{ht}^{(90)}}{P_{ht}^{(10)}} \right)$$

The **homogeneity ratio** is defined as:

$$H_{ht} = \frac{P_{ht}^{(10)}}{P_{ht}^{(90)}} = \exp(-\text{LogGap}_{ht})$$

where $H \in (0, 1]$ and higher values indicate tighter price clustering.

We compute two variants:

| Metric | Weighting | Interpretation |
|---|---|---|
| $H^{EQ}$ | Equal-weight across exporters | Each exporter is an independent price signal |
| $H^{VW}$ | Value-weighted by exporter trade | Reflects price distribution in actual trade flows |

### 1.4 Aggregation to HS6 Level

Year-level estimates are aggregated to HS6-level using value-coverage weighted medians:

$$\tilde{H}_h = \text{wtd.median}\left(\{H_{ht}\}, \{V_{ht} \cdot \text{Coverage}_{ht}\}\right)$$

This downweights years with poor UOM representation or low trade volume.

---

## 2. Divergence Testing

### 2.1 Motivation

If $H^{EQ} \neq H^{VW}$, large exporters occupy systematically different positions in the price distribution than small exporters. This divergence carries economic meaning:

- $\delta > 0$ ($H^{VW} < H^{EQ}$): High-value exporters in the tails (premium or discount positioning)
- $\delta < 0$ ($H^{VW} > H^{EQ}$): High-value exporters cluster near the median; small exporters are outliers

### 2.2 Test Statistic

Define:

$$\delta_{ht} = \text{LogGap}_{ht}^{VW} - \text{LogGap}_{ht}^{EQ}$$

Under the null hypothesis of no systematic price-value relationship, permuting exporter values $V$ while holding prices $P$ fixed should produce $\delta$ values centered at zero.

### 2.3 Permutation Inference

For each HS6-year with $n \geq 5$ exporters:

1. Compute observed $\delta_{ht}^{obs}$
2. Generate null distribution: for $r = 1, \ldots, R$, permute $\{V_e\}$ and compute $\delta^{(r)}$
3. Two-sided p-value: $p = \frac{1}{R} \sum_r \mathbf{1}\left[|\delta^{(r)}| \geq |\delta^{obs}|\right]$

Bootstrap resampling (with replacement) provides confidence intervals for $\delta^{obs}$.

### 2.4 HS6-Level Inference

Year-level p-values are combined using **Fisher's method**:

$$\chi^2 = -2 \sum_t \log(p_{ht})$$

which follows a $\chi^2_{2T}$ distribution under the null. Benjamini-Hochberg correction controls the false discovery rate across HS6 codes.

---

## 3. Rauch Classification

### 3.1 Conceptual Framework

Following Rauch (1999), goods are classified as:

| Category | Economic Meaning | Price Behavior |
|---|---|---|
| **Homogeneous** | Exchange-traded commodities | Tight clustering, law of one price |
| **Reference-priced** | Published benchmark prices | Moderate dispersion |
| **Differentiated** | Bilateral negotiation, quality variation | Wide price spread |

### 3.2 Classification Features

We classify HS6 codes using two features:

1. **Dispersion level** — $\widetilde{\text{LogGap}}_h$ (median across years)
2. **Temporal stability** — $\text{CV}_h = \frac{\text{sd}(\text{LogGap}_{ht})}{\text{mean}(\text{LogGap}_{ht})}$

The stability dimension captures the insight that truly homogeneous goods should exhibit not just low dispersion but *stable* dispersion over time.

### 3.3 Gaussian Mixture Model

We fit a GMM to the joint distribution of (LogGap, CV) across HS6 codes:

$$f(\mathbf{x}) = \sum_{k=1}^{K} \pi_k \, \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$$

Model selection (K = 2 or 3 components) uses BIC. Components are ordered by a homogeneity score:

$$S_k = 0.7 \cdot \mu_k^{\text{LogGap}} + 0.3 \cdot \mu_k^{\text{CV}}$$

The component with lowest $S_k$ is labeled "homogeneous," highest is "differentiated."

### 3.4 Stability Override

As a robustness check, HS6 codes classified as "homogeneous" but with CV above the 90th percentile are demoted to "reference-priced." This prevents volatile goods from receiving the homogeneous label.

### 3.5 Fallback Hierarchy

For small samples where GMM estimation is unreliable:

| Priority | Condition | Method |
|---|---|---|
| 1 | $n \geq 10$ | 2D GMM (dispersion + stability) |
| 2 | $n < 10$ | 1D GMM (dispersion only) |
| 3 | GMM fails | Quantile-based (33rd/67th percentiles) |

---

## 4. Rauch (1999) Benchmark Validation

### 4.1 Motivation

The GMM-based classification in Section 3 is purely data-driven. To validate this approach, we compare our classifications against Rauch's (1999) original taxonomy, which classified goods based on expert judgment and market structure analysis. This benchmark serves three purposes:

1. **Validation** — Confirms the methodology recovers known economic structure
2. **Calibration** — Identifies systematic biases or threshold issues
3. **Gap identification** — Highlights products where market character may have evolved since 1999

### 4.2 Original Rauch Classification

Rauch (1999) classified goods at the **SITC Revision 2, 4-digit** level (~1,189 codes) into three categories:

| Code | Category | Criteria |
|---|---|---|
| `w` | Homogeneous | Traded on organized exchanges |

| Code | Category | Criteria |
|---|---|---|
| `r` | Reference-priced | Prices published in trade journals |
| `n` | Differentiated | Neither of the above |

Two classification variants exist: **conservative** (stricter criteria for homogeneous/reference) and **liberal** (more inclusive). The conservative variant is used by default.

### 4.3 Concordance: SITC4 → HS6

Rauch's classifications predate modern HS nomenclature, requiring concordance from SITC Rev.2 to HS6. This mapping is inherently lossy:

| Challenge | Implication |
|---|---|
| Many-to-many mapping | One SITC4 may correspond to multiple HS6 codes with varying homogeneity |
| HS revisions | Multiple HS vintages (1992, 2002, 2007, 2012, 2017, 2022) require chained concordances |
| New products | ~15–25% of current HS6 codes have no SITC4 antecedent |
| Product evolution | Some goods have changed character since 1999 (e.g., electronics commoditization) |

The concordance is performed using the `concordance` R package (Liao et al.), which provides SITC–HS crosswalks.

### 4.4 Concordance Quality Tiers

We assess the reliability of each HS6's Rauch benchmark based on concordance quality:

| Tier | Criteria | Interpretation |
|---|---|---|
| **Tier 1 (Direct)** | Exact or near-exact SITC4 → HS6 match; unanimous Rauch category | High-confidence benchmark |
| **Tier 2 (Inherited)** | Good/acceptable concordance; unanimous Rauch category | Moderate confidence |
| **Tier 2 (Ambiguous)** | Multiple SITC4 sources disagree on Rauch category | Benchmark is modal category |
| **Tier 3 (Poor)** | Many-to-many mapping with poor concordance quality | Low confidence |
| **No Coverage** | HS6 has no SITC4 antecedent in Rauch data | Purely data-driven classification |

Concordance quality is determined by:

$$\text{Quality} = f(n_{\text{HS6 per SITC4}}, n_{\text{SITC4 per HS6}})$$

where exact matches (1:1) are highest quality, and dispersed mappings are lowest.

### 4.5 Aggregation to HS6

When an HS6 code maps to multiple SITC4 codes with different Rauch categories, we:

1. Take the **modal category** across all SITC4 sources
2. Flag `rauch_unanimous = FALSE` if sources disagree
3. Report the best concordance quality among sources

### 4.6 Benchmark Comparison Metrics

| Variable | Definition |
|---|---|
| `rauch_benchmark` | Original Rauch (1999) category via concordance |
| `rauch_unanimous` | TRUE if all SITC4 sources agree |
| `concordance_quality` | Best quality tier among SITC4 sources |
| `benchmark_tier` | Confidence tier (tier1_direct, tier2_inherited, etc.) |
| `benchmark_match` | TRUE if GMM classification equals Rauch benchmark |

### 4.7 Interpreting Agreement and Disagreement

**High agreement on Tier 1** (e.g., >70%): The GMM successfully recovers Rauch's expert classifications from price data alone—strong methodological validation.

**Systematic disagreements** warrant investigation:

| GMM Says | Rauch Says | Possible Explanations |
|---|---|---|
| Differentiated | Homogeneous | Quality differentiation emerged; market fragmentation |
| Homogeneous | Differentiated | Commoditization; standardization since 1999 |
| Reference | Homogeneous | Exchange trading declined; benchmark pricing replaced |

**Low coverage** (~20–40% of HS6 with no Rauch antecedent): Expected for modern product codes; the data-driven approach fills this gap.

### 4.8 Reconciliation to Final Classification

Rather than using the benchmark purely for validation, we reconcile the GMM and Rauch (1999) classifications into a single final category (`rauch_final`) using a **benchmark priority** rule:

| Condition | Final Source | Rationale |
|---|---|---|
| Tier 1 (direct match) | Rauch (1999) | High-confidence expert classification |
| Tier 2 (inherited) + unanimous | Rauch (1999) | Moderate confidence; expert judgment preferred |
| Tier 2 (ambiguous) | GMM | Conflicting expert sources; let data decide |
| Tier 3 (poor concordance) | GMM | Unreliable benchmark |

| Condition | Final Source | Rationale |
|---|---|---|
| No coverage | GMM | No expert classification available |

This approach leverages Rauch's expert judgment where it exists with confidence, while allowing the data-driven GMM to fill gaps and resolve ambiguities. The `final_source` column tracks whether each HS6's final classification came from `rauch_1999` or `gmm`.

---

## 5. Output Metrics

### 5.1 HS6-Level Summary

| Variable | Definition |
|---|---|
| `med_H_eq` | Median equal-weight homogeneity ratio |
| `med_H_vw` | Median value-weight homogeneity ratio |
| `cv_log_gap` | Coefficient of variation of LogGap across years |
| `rauch_category` | GMM classification: homogeneous / reference / differentiated |
| `rauch_original` | GMM classification before stability override |
| `posterior_prob` | GMM posterior probability for assigned category |
| `stability_flag` | TRUE if demoted due to high temporal instability |
| `med_delta` | Median divergence between VW and EQ measures |
| `flag_divergence` | TRUE if divergence is statistically significant (FDR-controlled) |
| `dominant_direction` | Direction of divergence: large_in_tails / large_in_center |
| `rauch_benchmark` | Original Rauch (1999) category (where available) |
| `benchmark_tier` | Concordance confidence: tier1_direct / tier2_inherited / no_coverage |
| `benchmark_match` | TRUE if GMM classification equals Rauch benchmark |
| `rauch_final` | **Reconciled classification** (Rauch 1999 where confident, else GMM) |
| `final_source` | Source of final classification: rauch_1999 / gmm |

### 5.2 Interpretation Guidelines

| med_H_eq | Interpretation |
|---|---|
| 0.8 – 1.0 | Highly homogeneous (commodity-like) |
| 0.5 – 0.8 | Moderate dispersion (reference-priced) |
| 0.2 – 0.5 | High dispersion (differentiated) |

| flag_divergence | dominant_direction | Interpretation |
|---|---|---|
| TRUE | large_in_tails | Large exporters at price extremes (quality tiers?) |
| TRUE | large_in_center | Large exporters at market price; small exporters are outliers |
| FALSE | — | No systematic price-value relationship |

---

## 6. Data Requirements

### 6.1 Input Fields

| Field | Type | Description |
|---|---|---|
| `year` | integer | Observation year |
| `exporter` | string | Exporter country/entity |
| `importer` | string | Importer country/entity |
| `hs6` | string | 6-digit HS code |
| `value` | numeric | Trade value |
| `quantity` | numeric | Trade quantity |
| `uom` | string | Unit of measure |

### 6.2 Sample Size Considerations

| Threshold | Default | Purpose |
|---|---|---|
| Minimum exporters per HS6-year | 10 | Reliable quantile estimation |
| Minimum coverage | 85% | Dominant UOM representativeness |
| Minimum years for HS6-level | 3 | Temporal stability estimation |
| Minimum exporters for divergence test | 5 | Permutation test validity |

---

## References


- Rauch, J.E. (1999). Networks versus markets in international trade. *Journal of International Economics*, 48(1), 7-35.
- Fraley, C. & Raftery, A.E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, 97(458), 611-631.
- Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate. *Journal of the Royal Statistical Society B*, 57(1), 289-300.
- Liao, S., In, S., Yotov, Y., & Zylkin, T. (2024). *concordance: Product Concordance*. R package. https://CRAN.R-project.org/package=concordance