

## **Spring 2025 Research Report**

### **Report Breakdown:**

1. Overview and Key Definitions
2. Data Structure and Methodology
3. Understanding the Data: Data Anomaly and Preliminary Screening
4. Data Cleaning: Multi-Winner Contracts and Value Splitting
5. Data Cleaning: Symbolic Contracts and Framework Agreements
6. Analysis: Vague CPVs Contracts
7. Analysis: Thresholds and Granularity
8. Analysis: Competitiveness Metrics and Sensitivity

## **1. Overview and Key Definitions**

The primary purpose of this preliminary research is to compare public procurement outcomes across 32 European (EU) countries and 44 Common Procurement Vocabulary (CPV) codes from 2018 to 2023 to measure fundamental differences in competitiveness. A key challenge in procurement analysis is that differences in contract concentration across countries may be driven by the type of goods/services being procured rather than true market competitiveness.

For instance, procurement of specialized projects (e.g., sports stadiums, defense contracts, etc.) typically involves fewer competing firms, due to the niche nature of these industries, whereas common procurement types (e.g., office supplies, IT services) naturally have higher competition due to lower barriers to entry. Simply put, if these differences are not accounted for, countries with specialized procurements will wrongly appear uncompetitive, when the issue is industry constraints, not market inefficiencies.

Thus, it is crucial to normalize each country's procurement competitiveness relative to the EU-wide competitiveness for each procurement type. This ensures we only flag countries as uncompetitive if they are truly underperforming compared to the expected competitiveness for their procurement mix.

This research introduces a multi-stage methodology grounded in three pillars: (1) classification of procurement categories as "Usual" or "Unusual" using the Pareto principle, (2)

assessment of market concentration via top-firm dominance in contract values, and (3) normalization against EU-wide CPV-specific benchmarks to control for sectoral differences in expected competitiveness.

As the research progressed, the scope expanded beyond competitiveness metrics to address the underlying structure and integrity of the data itself. This included identifying duplicate entries, contracts with implausible values (e.g. €10 trillion), multi-supplier frameworks reported as single awards, and widespread use of placeholder values (e.g. €0.01) that obscure actual financial commitments. Additionally, methods were developed to handle winner name segmentation, equal-value splitting in multi-award contracts, and pattern analysis for vague CPV usage and symbolic pricing behavior.

## **2. Data Structure and Methodology**

The dataset analyzed in this research originates from the TED (Tenders Electronic Daily) contract award notices. The study covers contract-level data from 2018 to 2023, comprising information on procurement across 32 European countries and spanning over 40 CPV (Common Procurement Vocabulary) codes at both 4-digit and 5-digit levels.

The raw dataset consists of approximately 6 million individual procurement contracts, each representing a unique award record. These entries vary in size, scope, and completeness, requiring careful structuring and filtering before any analysis can be conducted.

Each record includes the following core variables:

- ISO\_COUNTRY\_CODE: 2-letter code identifying the awarding country (e.g., DE for Germany).
- VALUE\_EURO: Total reported contract value in euros.
- CPV: A numeric code indicating the procurement type. These can appear at varying levels of granularity, typically 4-digit (e.g., 3310) or 5-digit (e.g., 33101).
- NUMBER\_OFFERS: Number of offers submitted for the contract.
- WIN\_NAME: Name(s) of the winning supplier(s).
- DT\_AWARD: Date the contract was awarded.

This “base structure” forms the cleaned and standardized table referred to as data\_filtered throughout the research. It serves as the analytical core of all subsequent competitiveness and structural analyses.

### **2.1) Employing the Pareto Principle to classify procurement as “usual” or “unusual”**

To distinguish between common and specialized procurement types, we apply the Pareto Principle (80/20 Rule) – a concept in procurement strategy, which suggests that a small subset of procurement categories accounts for the majority of contract awards (CIPS, 2023).

- "Usual" Procurement: The top 80% most frequent CPV categories, which typically attract higher competition due to their broad market availability.
- "Unusual" Procurement: The remaining 20% of procurement categories, often specialized industries with naturally fewer competitors (e.g., defense, aerospace).

This classification is a preliminary step to prevent countries with specialized procurement portfolios from being wrongly flagged as uncompetitive. Ideally, further steps to verify each procurement contract in each country will further solidify the understanding of the dynamic how each “usual” and “unusual” contracts function in each subregion.

### **2.2) Measuring procurement competitiveness within each country**

Competitiveness score is assessed by calculating the proportion of firms winning the top 75% of contract value within each country.

- If a few firms dominate most of the contract value, the market is concentrated and may indicate barriers to entry or inefficiencies.
- If contract values are more evenly distributed among firms, the market is more competitive.

After coding and statistical trials, I shifted the threshold from 50%, which was found to be too restrictive, leading to overly concentrated results. The 75% threshold provides a broader view of market concentration while still capturing competitiveness dynamics among the lower valued deals.

### **2.3) Adjusting competitiveness scores based on EU-wide procurement benchmarks**

Since different procurement types have different expected competition levels, we normalize each country's competitiveness relative to the EU-wide competitiveness score for each procurement category.

- Compute the EU-wide competitiveness score for each CPV type, which serves as a baseline for expected competition.
- Measure each country's competitiveness within each CPV category.
- Compare each country's competitiveness to the EU benchmark to determine if they are underperforming relative to expectations for their procurement mix.

By calculating the deviation from the EU-wide benchmark, we ensure that countries are only flagged as uncompetitive if they genuinely underperform, rather than simply procuring in naturally concentrated industries.

## **3. Understanding the Data: Data Anomaly and Preliminary Screening**

This section outlines the identification and treatment of key data anomalies in the TED procurement dataset, specifically focusing on extreme monetary values, implausible offer counts, and other inconsistencies that distort analytical outcomes such as competitiveness metrics and Lorenz curves.

### **3.1.1 Nature of Anomalies**

Several categories of extreme or inconsistent values were identified during initial data exploration:

- **Implausibly High Contract Values**
  - Contracts reporting values in excess of €10 trillion (e.g., “10T+”) were observed. These entries are highly unlikely to represent real procurement activities and are likely the result of input errors, placeholder values, or data integration issues.
- **Anomalous Offer Counts**
  - Numerous contracts recorded “999” offers—an implausibly high number relative to the dataset's norm (typically 1–10 offers). These are interpreted as symbolic values or errors rather than valid submissions.

- **Suspicious High-Value Contracts**

- Outliers such as a €9.9 billion contract for “Evaluation consultancy services” in France and a €10 billion contract for “Personal care products” in Latvia illustrate discrepancies between sector norms and reported figures.

- **Lot-Level Summation Inconsistencies**

- In some cases, the total value of master contracts did not reconcile with the sum of their associated lots or sub-contracts.

### **3.1.2 Cleaning Strategy and Filtering Process**

To mitigate the effects of the above anomalies, the following filtering procedures were applied:

- Exclusion of Symbolic or Erroneous Values: Contracts with clearly implausible values (e.g., >€1 trillion) or anomalous offer counts (e.g., 999) were filtered out.
- Temporary Removal of Vague CPVs: Contracts with non-informative CPV codes (e.g., 45000000) were set aside pending a refined classification strategy.

After applying these filters, the dataset was reduced from ~6 million contracts to approximately 1.2 million contracts. This refinement significantly improved the signal-to-noise ratio, allowing for more meaningful analysis of procurement patterns and competitiveness indicators.

## 4. Data Cleaning: Winner Segmentation and Equal Contract Splitting

In procurement datasets, many contracts involve multiple winning suppliers. However, these multi-winner contracts are inconsistently recorded—sometimes listing all suppliers in a single row using ambiguous delimiters (e.g., semicolons, slashes, commas), and other times spreading each supplier across separate rows. This inconsistency poses a significant obstacle for supplier-level analysis, particularly when evaluating market fragmentation, supplier dominance, or fairness in value distribution.

To address this, the analysis pursued a two-stage objective: first, to accurately segment composite supplier entries into individual entities; and second, to reallocate contract values fairly among all winners.

While initial attempts relied on Regular Expression-based string splitting, this method proved unreliable due to conflicts with legal naming conventions. Ultimately, a more robust solution was adopted—dividing total contract value equally across distinct supplier rows per contract—which offers greater accuracy, consistency, and analytical value across the dataset.

### 4.1 RegEx

A regular expression (RegEx) is used to split winner names based on a set of common delimiters:

$$\backslash s^*(?:;|\||/|---|,|\n|\t)+\backslash s^*$$

This pattern is designed to capture typical separators such as semicolons, vertical bars, slashes, dashes, commas, newlines, and tabs, while also trimming surrounding whitespace.

Each record with multiple winners is:

- Split into multiple rows (one per supplier)
- Assigned a proportion of the original contract value based on an equal division rule
- This segmentation is applied to a sample of data to validate coverage and assess edge cases in supplier naming conventions.
- This method did not work as I analyzed the contracts deeper.

#### **4.1.1 Limitations and Takeaways from Regex-Based Splitting**

Supplier names such as: Medtronic Danmark A/S, Asiana, spol. s r.o., “Contract management, a.s.---DS engineering PLUS, a.s.---Digital Construction Consulting s.r.o.---MANIFOLD GROUP s.r.o.” may be misinterpreted. In the second case, the comma in “Asiana, spol. s r.o.” could lead to an incorrect split into two entities. In the third case, the triple dash used as a separator does correctly indicate multiple firms, but if similar punctuation appears inside a name, it risks fragmenting a single legal entity.

This highlights a key limitation: the same delimiters used to separate suppliers are also present in many formal legal names. As a result, false segmentation can occur, leading to overestimated supplier counts and incorrect allocation of contract value.

Despite this limitation, classification and segmentation are typically not a problem for high-value contracts. High-value awards tend to report a smaller number of winners, potentially due to the nature of the contracts.

Though, the key issue with this that leads to the next method is the fact that for each contract, winners are reported on separated rows, not the same. Hence, this “within-row classification” method can be ruled out

#### **4.2 Dividing contract values**

In response to the failed method above, this method — dividing total contract value (VALUE\_EURO) by the count of associated supplier rows — provides a workable solution for handling multi-winner contracts. While it is susceptible to data entry issues such as missing supplier names, its performance on well-structured records confirms its utility for supplier-level financial analysis.

##### **4.2.1 Splitting by Suppliers Count**

To correct for this, a simple but effective method was applied: divide the total contract value equally among all suppliers listed under the same contract.

This allocation ensures that supplier-level totals are not overstated and provides a fair representation of how value is distributed in multi-winner contracts. The number of unique supplier rows per contract serves as a proxy for how many parties shared the award, forming the basis of the split.

##### **4.2.2 Validation Through Sampling**

To assess the method's reliability, three contract groups were sampled.

### (a) Contracts with the Highest Split Values

These correspond to contracts with large VALUE\_EURO and very few winners (e.g. 1–2). They represent cases with the largest financial weight per supplier.

**Findings:** Many of these showed VALUE\_EURO\_SPLIT == inf (infinity), due to supplier\_count == 0. Upon inspection, these contracts were missing WIN\_NAME entries entirely, suggesting either data quality issues or intentionally redacted information.

	ID_NOTICE	TED_NOTICE	ISO_COUNTRY	WIN_NAME	VALUE_EURO	CPV	NUMBER_OF	DT_AWARD	YEAR	supplier_count	VALUE_EURO_SPLIT
4513267	2022190770	ted.europa.e	FR		3500000	77211500		28/2/2022	2022	0	inf
5195763	2022682380	ted.europa.e	RO		641672.47	33690000			2022	0	inf
4273076	2021606734	ted.europa.e	FR		48375000	45231400		15/11/2021	2021	0	inf
1407212	2018354205	ted.europa.e	HU		543008.56	34631000	1	18/7/2018	2018	0	inf
5195765	2022682380	ted.europa.e	RO		641672.47	33690000			2022	0	inf
4273075	2021606734	ted.europa.e	FR		48375000	45231400		15/11/2021	2021	0	inf
2573609	2020115697	ted.europa.e	RO		3123402.23	33696500			2020	0	inf
5195764	2022682380	ted.europa.e	RO		641672.47	33690000			2022	0	inf
4273074	2021606734	ted.europa.e	FR		48375000	45231400		15/11/2021	2021	0	inf
5195762	2022682380	ted.europa.e	RO		641672.47	33690000			2022	0	inf
5195771	2022682380	ted.europa.e	RO		641672.47	33690000			2022	0	inf
5195761	2022682380	ted.europa.e	RO		641672.47	33690000			2022	0	inf
5195760	2022682380	ted.europa.e	RO		641672.47	33690000			2022	0	inf
5195759	2022682380	ted.europa.e	RO		641672.47	33690000			2022	0	inf
5195758	2022682380	ted.europa.e	RO		641672.47	33690000			2022	0	inf
5195757	2022682380	ted.europa.e	RO		641672.47	33690000			2022	0	inf
5195756	2022682380	ted.europa.e	RO		641672.47	33690000			2022	0	inf
5195755	2022682380	ted.europa.e	RO		641672.47	33690000			2022	0	inf
2573608	2020115697	ted.europa.e	RO		3123402.23	33696500			2020	0	inf
4273077	2021606734	ted.europa.e	FR		48375000	45231400		17/11/2021	2021	0	inf
5195766	2022682380	ted.europa.e	RO		641672.47	33690000			2022	0	inf
4273078	2021606734	ted.europa.e	FR		48375000	45231400		10/11/2021	2021	0	inf
1407272	2018354302	ted.europa.e	FR		454380	71730000		30/7/2018	2018	0	inf
4273105	2021606752	ted.europa.e	DK		52166207.2	45230000	2	11/10/2021	2021	0	inf
5195770	2022682380	ted.europa.e	RO		641672.47	33690000			2022	0	inf
2573606	2020115697	ted.europa.e	RO		3123402.23	33696500			2020	0	inf

Table 1: Equal Contract Splitting – Contracts with the Highest Values



## (b) Contracts with the Lowest Split Values

These involved contracts with symbolic values (e.g. 0, 0.01). After division, these resulted in VALUE\_EURO\_SPLIT = 0, even when there were valid suppliers.

**Findings:** These are typically framework agreements or placeholder entries where the full value is not intended to be disclosed. They require downstream filtering or separate treatment in competitiveness analysis.

	ID_NOTICE	CTED_NOTICE	ISO_COUNT	WIN_NAME	VALUE_EURO	CPV	NUMBER_OF	DT_AWARD	YEAR	supplier_cou	VALUE_EURO_SPLIT
4194552	2021559504	ted.europa.e	RO	DELOITTE CC	0	71319000	1	24/5/2021	2021	1	0
5589724	2023329181	ted.europa.e	RO	ENEL ENERG	0	9310000	2	18/4/2023	2023	1	0
4067273	2021482310	ted.europa.e	RO	EDITRONIC I	0	33124200	1	3/8/2021	2021	1	0
6067244	2023667848	ted.europa.e	CZ		0	77341000			2023	2	0
6067243	2023667848	ted.europa.e	CZ		0	77341000			2023	2	0
6067242	2023667848	ted.europa.e	CZ	VALCANO a.s	0	77341000	10	9/3/2022	2023	2	0
6067241	2023667848	ted.europa.e	CZ	ROCKNET s.r	0	77341000	10	9/2/2022	2023	2	0
2943821	2020340066	ted.europa.e	CZ	Ha-vel intern	0	64200000	1	30/6/2020	2020	1	0
3489456	2021136619	ted.europa.e	DK		0	33000000			2021	33	0
5897111	2023545160	ted.europa.e	CZ	zrušeno # ve	0	79710000	1	6/9/2023	2023	1	0
3489455	2021136619	ted.europa.e	DK		0	33000000			2021	33	0
3489453	2021136619	ted.europa.e	DK	Vingmed A/S	0	33000000	6	20/1/2021	2021	33	0
3489452	2021136619	ted.europa.e	DK	Medtronic D	0	33000000	6	20/1/2021	2021	33	0
3489451	2021136619	ted.europa.e	DK	Boston Scien	0	33000000	6	20/1/2021	2021	33	0
3489450	2021136619	ted.europa.e	DK	Merit Medica	0	33000000	3	20/1/2021	2021	33	0
3489449	2021136619	ted.europa.e	DK	Medtronic D	0	33000000	3	20/1/2021	2021	33	0
3489448	2021136619	ted.europa.e	DK	Boston Scien	0	33000000	3	20/1/2021	2021	33	0
3489447	2021136619	ted.europa.e	DK	Abbott Medic	0	33000000	7	20/1/2021	2021	33	0
3489446	2021136619	ted.europa.e	DK	Medtronic D	0	33000000	7	20/1/2021	2021	33	0
3489454	2021136619	ted.europa.e	DK		0	33000000			2021	33	0
6067858	2023668277	ted.europa.e	CZ	DATRON, a.s	0	48000000	1	19/7/2023	2023	1	0
830770	202244161	ted.europa.e	CZ	Contract ma	0	71000000	4	31/12/2021	2022	1	0
6071896	2023670435	ted.europa.e	CZ	Crayon Cze	0	48000000	1	18/8/2023	2023	1	0
277101	201920444	ted.europa.e	MK	ESI temporar	0	79620000	6	26/12/2018	2019	1	0
220967	201872104	ted.europa.e	CZ	fãesky o po	0	79800000	2	20/12/2017	2018	1	0
35678	20194445	ted.europa.e	NO	Entelios AS	0	9310000	5	19/12/2018	2019	1	0
88850	20222044	ted.europa.e	NO	Trollfjord Kra	0	9310000	3	21/12/2021	2022	1	0

Table 2: Equal Contract Splitting – Contracts with the Lowest Values

### (c) Middle-of-Range Contracts (Unsorted Sample)

A random subset of contracts was also inspected without sorting, to observe typical behavior.

**Findings:** The splitting method functioned as intended. Supplier names were consistently logged, and value distribution per supplier was reasonable and numerically consistent with expectations.

	ID_NOTICE	(TED_NOTICE	ISO_COUNT	WIN_NAME	VALUE_EURO	CPV	NUMBER_OF	DT_AWARD	YEAR	supplier_cou	VALUE_EURO_SPLIT
54894	20204400	ted.europa.e	PL	Farmacol-Lo	884047.42	33600000	4	9/12/2019	2020	81	10914.1657
54895	20204400	ted.europa.e	PL		884047.42	33600000			2020	81	10914.1657
54896	20204400	ted.europa.e	PL	Centrala farr	884047.42	33600000	2	9/12/2019	2020	81	10914.1657
54897	20204400	ted.europa.e	PL	NEUCAS.A.	884047.42	33600000	4	9/12/2019	2020	81	10914.1657
54898	20204400	ted.europa.e	PL	ASPEN PHAR	884047.42	33600000	2	9/12/2019	2020	81	10914.1657
54899	20204400	ted.europa.e	PL	NEUCAS.A.	884047.42	33600000	4	9/12/2019	2020	81	10914.1657
54900	20204400	ted.europa.e	PL	Farmacol-Lo	884047.42	33600000	3	9/12/2019	2020	81	10914.1657
54901	20204400	ted.europa.e	PL	PGF S.A.---Uf	884047.42	33600000	4	9/12/2019	2020	81	10914.1657
54902	20204400	ted.europa.e	PL	Salus Interna	884047.42	33600000	4	9/12/2019	2020	81	10914.1657
54903	20204400	ted.europa.e	PL	Aesculap Chi	884047.42	33600000	1	9/12/2019	2020	81	10914.1657
54904	20204400	ted.europa.e	PL	PGF S.A.---Uf	884047.42	33600000	4	9/12/2019	2020	81	10914.1657
54905	20204400	ted.europa.e	PL	ASCLEPIOS S	884047.42	33600000	5	9/12/2019	2020	81	10914.1657
54906	20204400	ted.europa.e	PL		884047.42	33600000			2020	81	10914.1657
54907	20204400	ted.europa.e	PL	Salus Interna	884047.42	33600000	3	9/12/2019	2020	81	10914.1657
54908	20204400	ted.europa.e	PL	Salus Interna	884047.42	33600000	4	9/12/2019	2020	81	10914.1657
54909	20204400	ted.europa.e	PL	Delfarma Sp.	884047.42	33600000	2	9/12/2019	2020	81	10914.1657
54910	20204400	ted.europa.e	PL		884047.42	33600000			2020	81	10914.1657
54911	20204400	ted.europa.e	PL	PGF S.A.---Uf	884047.42	33600000	2	9/12/2019	2020	81	10914.1657
54912	20204400	ted.europa.e	PL	Salus Interna	884047.42	33600000	4	9/12/2019	2020	81	10914.1657
54913	20204400	ted.europa.e	PL	PGF S.A.---Uf	884047.42	33600000	3	9/12/2019	2020	81	10914.1657
54914	20204400	ted.europa.e	PL		884047.42	33600000			2020	81	10914.1657
54915	20204400	ted.europa.e	PL	OPTIFARMA s	884047.42	33600000	2	9/12/2019	2020	81	10914.1657
54916	20204400	ted.europa.e	PL	Salus Interna	884047.42	33600000	3	9/12/2019	2020	81	10914.1657
54917	20204400	ted.europa.e	PL	Farmacol-Lo	884047.42	33600000	3	9/12/2019	2020	81	10914.1657
54918	20204400	ted.europa.e	PL	PGF S.A.---Uf	884047.42	33600000	2	9/12/2019	2020	81	10914.1657
54919	20204400	ted.europa.e	PL	Salus Interna	884047.42	33600000	3	9/12/2019	2020	81	10914.1657
54920	20204400	ted.europa.e	PL	NEUCAS.A.	884047.42	33600000	3	9/12/2019	2020	81	10914.1657
54921	20204400	ted.europa.e	PL	PGF S.A.---Uf	884047.42	33600000	2	9/12/2019	2020	81	10914.1657
54922	20204400	ted.europa.e	PL	Salus Interna	884047.42	33600000	3	9/12/2019	2020	81	10914.1657
54923	20204400	ted.europa.e	PL	NEUCAS.A.	884047.42	33600000	4	9/12/2019	2020	81	10914.1657
54924	20204400	ted.europa.e	PL		884047.42	33600000			2020	81	10914.1657

Table 3: Equal Contract Splitting – Middle-Range, Random Sample Contracts

### Key Takeaway

This new approach is both scalable and interpretable. While it will require flagging and filtering edge cases (missing names, symbolic values), it accurately reflects supplier participation in most contracts and forms a solid foundation for deeper analysis of market share, competition, and fragmentation.

## **5. Data Cleaning: Low-Value, Symbolic Contracts and Framework Agreements**

A subset of contracts in the dataset reported extremely low or near-zero total values, often under 1 euro. These entries warrant closer inspection due to their potential impact on value-based metrics and fairness indicators.

This section investigates the nature of these low-value contracts to determine whether they represent:

- Symbolic or placeholder reporting (e.g., 0.01 EUR)
- Framework agreements involving multiple pre-approved suppliers
- Patterns of usage concentrated in specific countries, buyers, or procurement sectors (CPV codes)

### **Case 1 Review: Denmark, Contract ID 2021136619**

An example of this behavior is observed in a Danish contract published under ID [2021136619](#). The notice includes 30+ separate contract awards, each with a reported value of 0.01 DKK and linked to distinct suppliers across multiple lots. Upon reviewing the contract notice, the following features were observed:

- The contract is divided into 33 lots, each assigned to multiple suppliers
- Multiple suppliers are listed under the same lot, each receiving a duplicate 0.01 DKK value
- The contract titles and CPV codes pertain to medical equipment and devices

The award structure and repeated low values suggest a potential framework agreement where actual purchases occur later via call-offs.

## **Case 2 Review: Germany, Contract ID [216865-2023](#)**

This contract is divided into multiple lots, with Lot 1 alone awarded to over 30 suppliers, each listed under separate contract numbers. (I utilized ChatGPT for this, the list goes on forever inside the TED contract reports)

### **1. Symbolic or Placeholder Values**

- Each supplier received a contract with a reported value of 0.01 EUR, despite the initial estimated value being 730.00 EUR.
- The identical, minimal values across all contracts strongly indicate symbolic pricing, used to comply with publication requirements rather than to reflect actual transaction amounts.

### **2. Framework Agreement Structure**

- The contract appears to be a framework agreement (FA), where:
  - Multiple suppliers are pre-qualified.
  - Actual services and payments occur through future, on-demand call-offs.
  - **The minimal value in TED reflects no immediate financial commitment.**

### **3. Multiple Individual Contracts for a Single Lot**

- Lot 1 was awarded to 30+ distinct suppliers under separate contract numbers.
- This mirrors the standard FA model, where multiple suppliers are awarded in parallel to ensure availability and competitive flexibility over time.

### **4. Consistency Across Data Points**

- All contracts share the same award date (13/03/2023).
- CPV codes and project descriptions align with general IT consulting and support activities.
- Suppliers include a mix of SMEs and large firms, suggesting open competition and a broad vendor base.

As part of the deeper investigation into low-value contracts, I reviewed four specific public procurement notices where contracts were awarded at extremely low reported values (ranging from 0.01 EUR to a few euros). These cases reveal important nuances behind "symbolic" reporting practices and help frame how such contracts should be treated in the analysis.

### Cases Reviewed for <1 Euros:

#### 1. [Netherlands](#) – Gemeente Rotterdam: Environmental Advisory Services

- **Entity:** Gemeente Rotterdam
- **CPV:** 79620000 (Supply services of personnel, including temporary staff)
- **Summary:**
  - The initial published contract value was 0.01 EUR.
  - Multiple small follow-up lots were awarded to the same supplier (MilieuProfiel BV), each valued between approximately 46–63 EUR.
  - This setup reflects a **framework agreement**, where the symbolic low value marked a placeholder. Actual micro-contracts were issued later for specific advisory services.
- **Key dynamic:** Standard framework setup with symbolic reporting for compliance.

#### 2. [France](#) – Lamballe Communal Procurement for Medical Supplies

- **Entity:** Service Commun d'Achats, Lamballe
- **CPV:** 33100000 (Medical equipment) and related codes
- **Summary:**
  - Numerous lots for incontinence products, hygiene equipment, and medical devices were reported at 1.00 EUR per supplier.
  - A multi-supplier **framework agreement** was established, with the real financial flows occurring only when call-offs were issued post-framework setup.
  - **Key dynamic:** Pre-qualification of multiple suppliers without upfront financial commitment; symbolic reporting for transparency.

### 3. [Latvia](#) – Rural Advisory Center Bus Rentals

- **Entity:** Latvian Rural Advisory and Training Centre
- **CPV:** 60100000 (Road transport services)
- **Summary:**
  - Contracts were awarded across many towns with reported values between 0.90 EUR and 5.50 EUR.
  - The purpose was to set up a network for rural bus services under EU-funded projects, with actual service orders happening later.
- **Key dynamic:** Symbolic contracts to establish eligibility for small local operators, particularly supporting rural transport projects.

### 4. [Germany](#) – Max Planck Institute for Psychiatry: Database Systems

- **Entity:** Max-Planck-Institut für Psychiatrie, Munich
- **CPV:** 48610000 (Database systems)
- **Summary:**
  - The reported value for each awarded lot was 0.01 EUR.
  - This reflects a research-related procurement where actual commercial terms were handled confidentially outside the published notice.
- **Key dynamic:** Placeholder values used where confidentiality clauses or later commercial negotiation govern real financial transactions.

### Key Takeaway from <1 Euro Contracts:

1. **Framework Agreements Distort Competitiveness Metrics:**

Many low-value contract records result from framework setups, not from true open-market competition events. These need to be treated separately in competitiveness or corruption analysis.
2. **Multi-Supplier Awards and Micro-Lots are Indicators:**

Symbolic frameworks often involve awarding many suppliers simultaneously, each with very low nominal amounts, especially in sectors like medical supplies, consulting services, or transport.
3. **Real Financial Activity Happens After Publication:**

True spending flows (and thus true competitive dynamics) occur only during the execution phase (e.g., call-offs), which are typically not captured within the initial notice.
4. **Different Countries, Same Patterns:**

Across the Netherlands, France, Latvia, and Germany, although procedural differences

exist, symbolic low-value contracts serve similar administrative functions rather than indicating actual economic transactions.

## Cases Reviewed for >1 Euros:

### 1. [United Kingdom](#) – NHS Shared Business Services: Facilities Management Framework

- **Entity:** NHS Shared Business Services Ltd (NHS SBS)
- **CPV:** 79993100 (Facilities management services)
- **Summary:**
  - This was a **50-lot framework** covering a wide range of facilities management services (building management systems, plumbing, locksmith services, HVAC maintenance, etc.).
  - The **reported lowest contract value** was **1.00 GBP**, and the **highest offer** was **500 million GBP** across all lots.
  - The framework is designed for **4 years** and allows public sector bodies to call off specific services under agreed terms.
- **Key dynamic:** The symbolic value (1 GBP) again reflects **pre-qualification** onto a framework, where real financial commitments occur later through service call-offs based on actual needs.

### 2. [United Kingdom](#) – NHS Shared Business Services: Urology Products

- **Entity:** NHS Shared Business Services Ltd (NHS SBS)
- **CPV:** 33141200 (Catheters and related urology supplies)
- **Summary:**
  - A **multi-lot framework** providing access to a wide range of urology products (catheters, drainage bags, support services).
  - Contract awards showed a **lowest offer of 1.00 GBP** and a **highest offer up to 118 million GBP**.
  - Multiple suppliers (e.g., Bard Ltd, Clinisupplies, ConvaTec, Flexicare) were awarded across different lots.
- **Key dynamic:** Similar to previous symbolic cases, the 1 GBP figure serves as a **technical placeholder**. Real procurement happens dynamically over the framework's life, based on actual orders placed by healthcare institutions.

### 3. [Germany](#) – BARMER Health Insurance Fund: Generic Pharmaceutical Framework (Tranche 11a 2018)

- **Entity:** BARMER Justiziariat-Vergabestelle
- **CPV:** 33600000 (Pharmaceutical products), **most common procurement types**
- **Summary:**
  - Framework agreement covering discounts for a wide range of generic pharmaceuticals (multiple lots including Abacavir, Amantadin, Anagrelid, etc.).
  - 46 Suppliers
  - Total reported value: **1.01 EUR**.
  - Contracts awarded were structured as non-exclusive rebate agreements
  - No minimum purchase quantities guaranteed; actual sales volumes depend on prescriptions and patient usage.
- **Key dynamic:**
  - Symbolic low-value contract formalizing rebate conditions without financial commitments at award stage.

### 4. [Germany](#) – AOK PLUS: Non-Exclusive Discount Agreements (Leuprorelin and Posaconazol)

- **Entity:** AOK PLUS – Die Gesundheitskasse für Sachsen und Thüringen
- **CPV:** 33600000 (Pharmaceutical products)
- **Summary:**
  - Non-exclusive rebate agreements for two active substances (Leuprorelin and Posaconazol).
  - Total reported value: **6.00 EUR**.
  - "Open-house" model: all qualifying pharmaceutical companies could join under pre-set conditions without individual negotiation.
  - No minimum guaranteed volume; actual utilization dependent on prescriptions.
- **Key dynamic:**
  - Symbolic contract value marking eligibility for rebates, not tied to any immediate transaction.



## 5. [Germany](#) – AOK PLUS: Non-Exclusive Discount Agreements (Teriparatid and Raltegravir)

- **Entity:** AOK PLUS – Die Gesundheitskasse für Sachsen und Thüringen
- **CPV:** 33600000 (Pharmaceutical products)
- **Summary:**
  - Similar "open-house" non-exclusive rebate agreements for the pharmaceutical products Teriparatid and Raltegravir.
  - Total reported value: **3.00 EUR**.
  - Agreements available to any qualifying pharmaceutical company under fixed terms during a set timeframe.
- **Key dynamic:**
  - Symbolic reported value; real contractual and financial impact occurs after individual prescriptions trigger rebates.

## 6. [Germany](#) – AOK PLUS: Non-Exclusive Discount Agreement for Octreotid

- **Entity:** AOK PLUS – Die Gesundheitskasse für Sachsen und Thüringen
- **CPV:** 33600000 (Pharmaceutical products)
- **Summary:**
  - Non-exclusive framework agreement for Octreotid.
  - Total reported value: **3.00 EUR**.
  - Operated under a similar "open-house" model without exclusivity or guaranteed sales volume.
- **Key dynamic:**
  - Symbolic placeholder value linked to administrative setup of rebate rights under statutory health insurance law.

## **6. Analysis: Vague CPVs Contracts**

Vague CPV (Common Procurement Vocabulary) codes refer to entries that use broad or overly general classifications, typically indicated by CPV codes ending in multiple trailing zeros (e.g., 45000000 for “Construction work”). These codes lack the granularity necessary to pinpoint specific goods or services and often reflect either low reporting precision or intentional opacity in procurement documentation.

### **Methodology:**

#### **Value Tier Classification:**

- The dataset was split into Top 20% and Bottom 80% based on contract value (VALUE\_EURO).
- Vague contracts were identified using CPV codes ending in multiple trailing zeros.

#### **Two Scenarios Analyzed:**

- Raw Data: The original dataset, which includes potential duplicates.
- No-Duplicate Dataset: A cleaned version excluding repeated entries based on contract ID and URL.

### **Metric:**

- For both datasets, the percentage of vague contracts was calculated within each value tier.

### **Findings**

The assumption that vague contracts might be concentrated in low-value procurements or strategically obscure high-value transactions is not clearly supported by the data. In fact, vague contracts are relatively evenly distributed across both value tiers, with only minor differences between them.

### **Supporting Evidence (No-Duplicate Data):**

#### **Bottom 80%:**

607,770 vague contracts out of 1,149,132 total → 52.89%

#### **Top 20%:**

145,981 vague contracts out of 288,973 total → 50.52%

## Raw Dataset Comparison (With Duplicates):

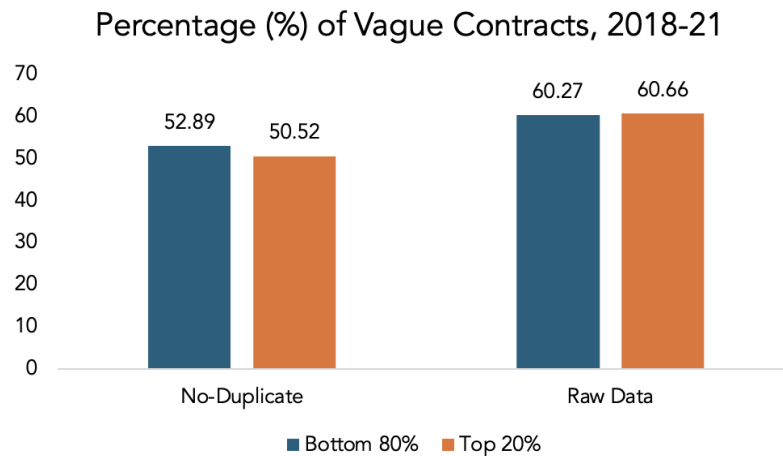
**Bottom 80%:** 60.27% vague (2,712,956 / 4,501,057)

**Top 20%:** 60.66% vague (682,683 / 1,125,347)

This suggests that including duplicate rows artificially inflates the appearance of vagueness across the board—particularly in high-value contracts. But even after removing duplicates, the proportion remains relatively high and stable.

### Interpretation:

- Vague CPVs are widespread, not just limited to low-value or low-risk contracts. This challenges the narrative that vagueness is a tool for obfuscating only large procurements.
- The small difference between Top 20% and Bottom 80% in both datasets (approx. 2.4 percentage points in no-duplicate data) suggests vagueness is a systemic reporting norm, not a value-tier-specific tactic.
- Duplicate rows may exaggerate trends, especially in large frameworks with multiple awards, but the trend remains present even when controlling for this.



No-Dupl.	Vague	Total	Percentage of Vague (%)
Bottom 80%	607770	1149132	52.88948528
Top 20%	145981	288973	50.51717635

Raw	Vague	Total	Percentage of Vague (%)
Bottom 80%	2712956	4501057	60.2737535
Top 20%	682683	1125347	60.6642218

Figure 1: Percentage (%) of Vague Contracts at Bottom 80% and Top 20%

Figure 2: Table Displaying Percentage (%) and Ratio of Vague Contracts

## By-Country Analyses for Vague Contracts

### Common Top 5 Countries Across Both Datasets

- Finland (FI)
- Slovenia (SI)
- North Macedonia (MK)
- Denmark (DK)
- (Only IS appears in Raw; NL appears in No-Dupl.)

These countries consistently rank among the top, regardless of whether duplicates are removed. This may imply systematic, or incidental, use of vague CPVs in procurement reporting.

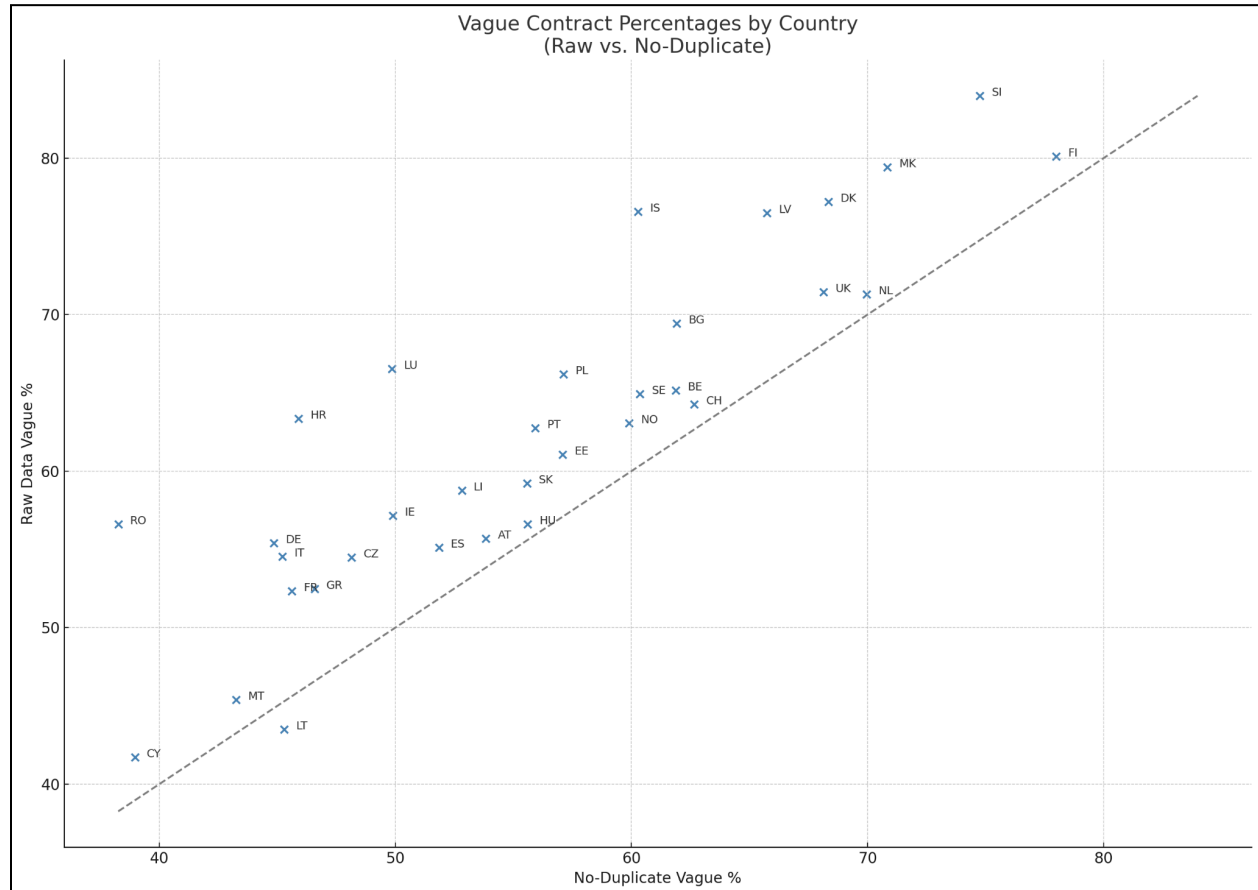


Figure 3: Vague Contract Percentages by Country (Raw and No-Duplicate)

Percentage (%) of Vague Contracts by Country (Raw Data)			
COUNTRY	TOTAL	VAGUE	VAGUE %
SI	242769	203866	83.98
FI	55314	44308	80.10
MK	11751	9334	79.43
DK	51330	39627	77.20
IS	2455	1880	76.58
LV	87237	66739	76.50
UK	141549	101116	71.44
NL	58382	41621	71.29
BG	150730	104650	69.43
LU	8393	5585	66.54
PL	1011466	669624	66.20
BE	49062	31967	65.16
SE	116069	75371	64.94
CH	17854	11472	64.25
HR	79973	50674	63.36
NO	34016	21451	63.06
PT	60063	37681	62.74
EE	40896	24973	61.06
SK	30091	17816	59.21
LI	359	211	58.77
IE	26813	15322	57.14
RO	1133832	641937	56.62
HU	69339	39248	56.60
AT	32796	18260	55.68
DE	452679	250744	55.39
ES	333480	183816	55.12
IT	182694	99642	54.54
CZ	202610	110387	54.48
GR	60020	31498	52.48
FR	690467	361330	52.33
MT	4804	2181	45.40
LT	182107	79220	43.50
CY	5004	2088	41.73

Table 4: Percentage (%) of Vague Contracts by Country, Raw Data

Percentage (%) of Vague Contracts by Country (No-duplicate Data)			
COUNTRY	TOTAL	VAGUE	VAGUE %
FI	21655	16888	77.99
SI	31523	23562	74.75
MK	1910	1353	70.84
NL	38319	26811	69.97
DK	13957	9538	68.34
UK	65906	44906	68.14
LV	17027	11194	65.74
CH	14249	8928	62.66
BG	55517	34371	61.91
BE	20976	12979	61.88
SE	41850	25261	60.36
IS	997	601	60.28
NO	18189	10894	59.89
PL	162121	92608	57.12
EE	8638	4931	57.08
PT	20598	11520	55.93
HU	19277	10719	55.61
SK	10518	5845	55.57
AT	18567	9996	53.84
LI	231	122	52.81
ES	108892	56455	51.84
IE	7269	3627	49.90
LU	4017	2003	49.86
CZ	78602	37840	48.14
GR	18043	8404	46.58
HR	15105	6933	45.90
FR	202468	92360	45.62
LT	22517	10199	45.29
IT	57831	26147	45.21
DE	242398	108734	44.86
MT	3028	1310	43.26
CY	2266	883	38.97
RO	93644	35829	38.26

Table 5: Percentage (%) of Vague Contracts by Country, No-Duplicate Data

## 7. Analysis: Thresholds and Granularity

I specifically analyzed the impact of CPV granularity (4-digit vs. 5-digit classification) and threshold splits (70/30 vs. 80/20) on competitiveness scores. The analysis is conducted separately for Usual and Unusual procurements to assess their stability across different configurations.

Initial findings showed that grouping “Usual” and “Unusual” procurements together in the analyses resulted in a nearly-identical competitiveness score. Hence, the approach was taken to deepen the analyses further.

Note: the data used to analyze this portion of the report has not been processed

### 7.1 Key Findings

#### 7.1.1 CPV Granularity: 4-Digit and. 5-Digit

This comparison evaluates whether increasing CPV classification detail (moving from 4-digit to 5-digit) significantly alters competitiveness scores.

#### Key Findings:

- **Usual Procurements:** Scatterplots indicate a strong linear relationship between 4-digit and 5-digit CPV scores, suggesting that increased CPV granularity does not significantly change competitiveness rankings. The line of best fit (LOBF) appears to be close to 1, indicating that using 4-digit CPVs is sufficient for Usual procurements.
- **Unusual Procurements:** The relationship is still relatively strong but shows more scatter compared to Usual procurements. This suggests that some Unusual procurements are affected by CPV granularity, but the overall effect is not drastic.
- **Conclusion:** CPV granularity does not meaningfully alter competitiveness scores for Usual procurements, but Unusual procurements exhibit some sensitivity.

### 2.1.1 Visualizations

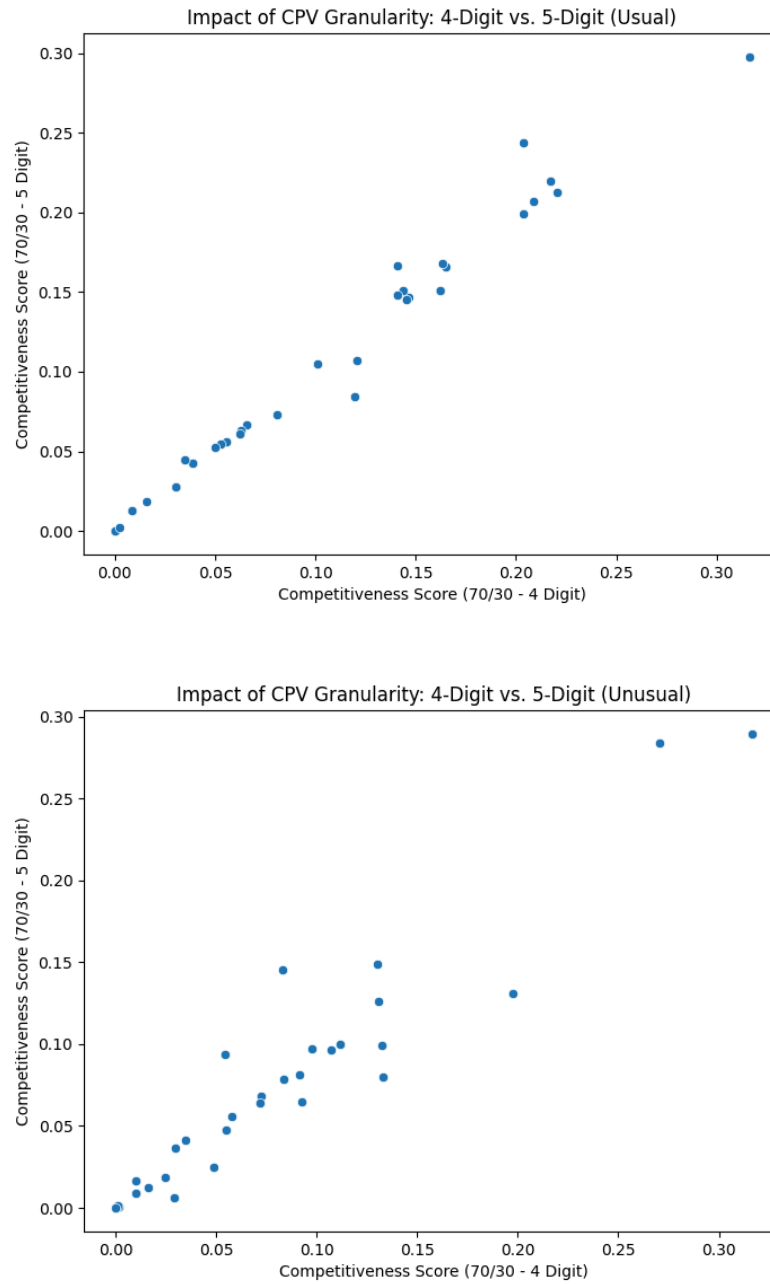


Figure 4 and 5: Granularity Comparison for Usual and Unusual Procurements within the 70-30 Threshold, at 4 and 5-digits Level



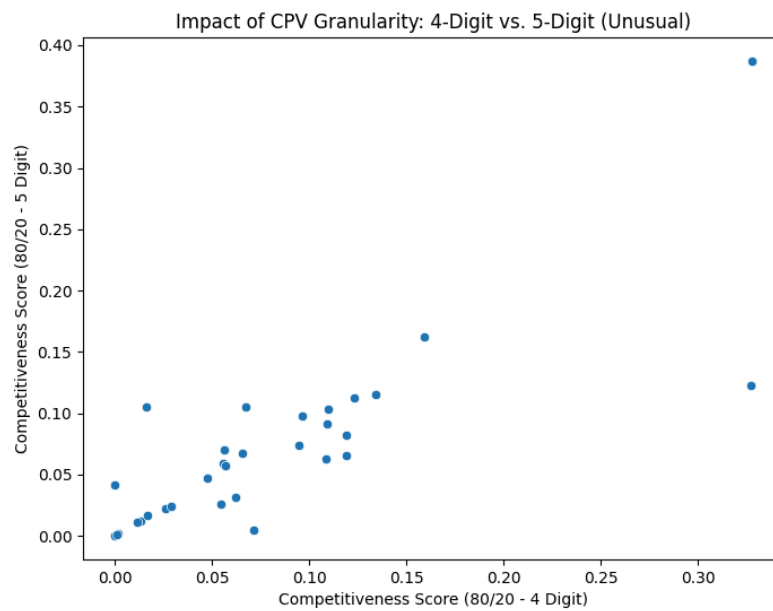
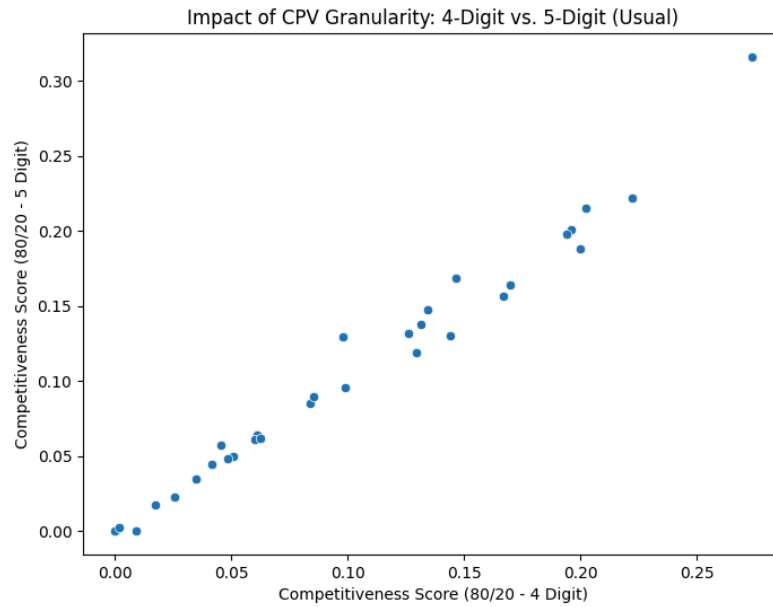


Figure 6 and 7: Granularity Comparison for Usual and Unusual Procurements within the 80-20 Threshold, at 4 and 5-digits Level

### 7.1.2 Threshold Impact: 70/30 vs. 80/20

This comparison assesses whether shifting the threshold for defining Usual and Unusual procurements (from 70/30 to 80/20) changes competitiveness rankings.

#### Key Findings:

- **Usual Procurements:** Scatterplots demonstrate a strong linear correlation between the two threshold splits. This suggests that competitiveness rankings remain stable regardless of whether 70/30 or 80/20 is used.
- **Unusual Procurements:** There is more scatter compared to Usual procurements, indicating that Unusual categories are more sensitive to threshold changes. This suggests that changing the threshold classification can alter competitiveness rankings for Unusual procurements.
- **Conclusion:** Usual procurements remain stable across threshold splits, while Unusual procurements are slightly more affected by changes in threshold selection.

## 2.1.2 Visualizations

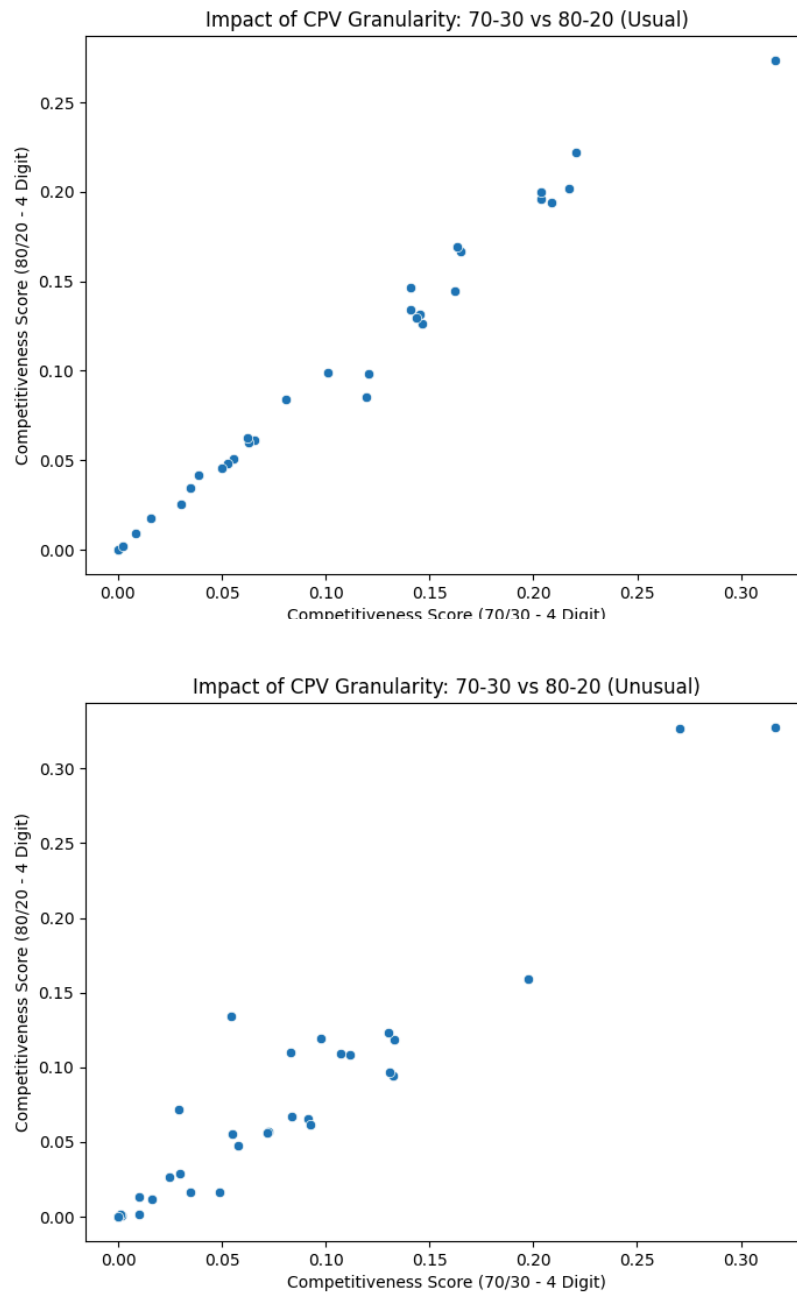


Figure 8 and 9: Granularity Comparison for Usual and Unusual Procurements between the 70-30 and 80-20 Threshold, at 4-digits Level

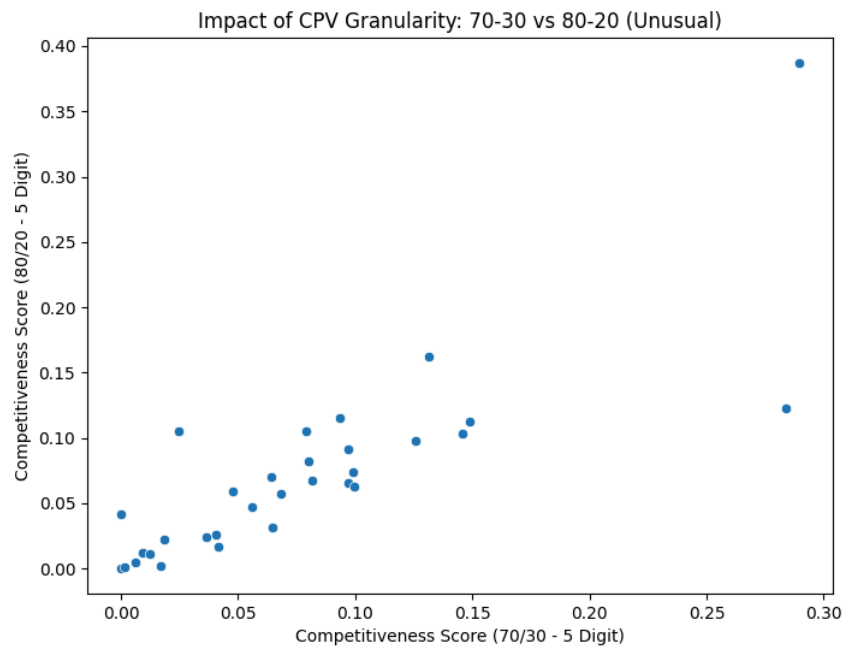
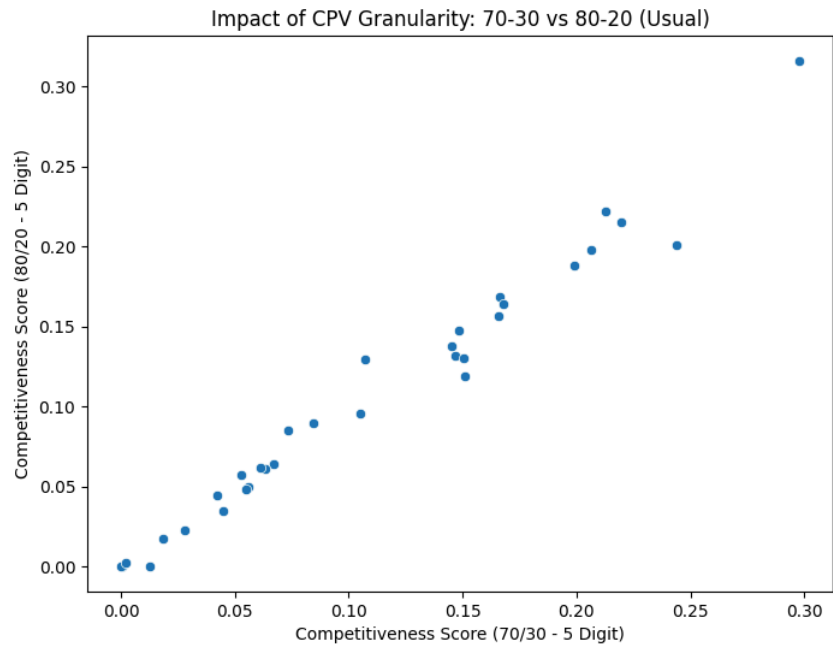


Figure 10 and 11: Granularity Comparison for Usual and Unusual Procurements between the 70-30 and 80-20 Threshold, at 5-digits Level

### **7.1.3. Initial Takeaways**

1. CPV granularity (4-digit vs. 5-digit) does not significantly impact competitiveness rankings for Usual procurements.
2. Threshold selection (70/30 vs. 80/20) does not affect Usual procurements but has a noticeable impact on Unusual procurements.
  - This suggests that the competitiveness of Unusual industries is more sensitive to the definition of usual vs. unusual procurement types.