Which Industries Are Actually the Defence Industries? AI Meets IO

Esteban Fernández Vázquez, Víctor Llaneza Menéndez, André Carrascal-Incera
Universidad de Oviedo

Abstract

The question of which industries constitute the defence industrial base resists straightforward answers within standard statistical frameworks. Military production is distributed across fabricated metals, electronics, aerospace, chemicals, and other manufacturing sectors that are not exclusively or visibly defence-oriented. This paper proposes a two-step methodology that bridges natural language processing and input-output analysis to identify and quantify these linkages at the country level across EU member states. In the first step, we compute semantic similarity scores between the 22 equipment categories of the Wassenaar Arrangement Munitions List and the 64 product groups of the Classification of Products by Activity (CPA 2.1) used in Eurostat's FIGARO supply tables, using sentence-embedding models. In the second step, this universal similarity matrix is combined with country-specific supply tables to produce a defence readiness profile φ(m, i, c) for each EU member state — quantifying how strongly each military equipment category is linked to each NACE industry through the country's actual production structure. We then apply this framework to simulate the sectoral production impacts of the EU Readiness 2030 strategy, operationalised as each member state closing its gap to the NATO 2% GDP defence spending target. Results show significant cross-country heterogeneity in both the total production burden and its sectoral distribution, driven jointly by differences in defence spending gaps and in national supply table structures. The framework provides a principled, reproducible tool for translating defence investment scenarios into disaggregated sectoral demand vectors.

Keywords: defence economics, input-output analysis, semantic similarity, FIGARO, EU Readiness 2030, Munitions List, NACE, supply tables

JEL codes: C67, H56, L64, O25, F52

1. Introduction

The war in Ukraine and the subsequent European rearmament debate have placed the defence industrial base at the centre of economic policy discussions in ways not seen since the Cold War. The European Commission's ReArm Europe initiative and the associated EU Readiness 2030 strategy have set ambitious targets for increasing defence spending across member states, with the NATO 2% of GDP benchmark serving as the primary operational anchor. Yet despite the urgency of these commitments, the economic infrastructure needed to assess their sectoral and regional consequences remains underdeveloped.

The fundamental difficulty is taxonomic. Standard industrial classifications — in particular NACE Rev. 2 — do not include a dedicated defence sector. Military production is distributed across a range of industries: weapons and ammunition appear partly in NACE C25 (fabricated metal products), military vehicles in C29 and C30, aircraft in C30, electronics and communications equipment in C26, energetic materials in C20. The boundaries are blurred further by dual-use production, where the same establishments, machines, and workers produce goods for both civilian and military markets. This distributional character of military production means that aggregate defence spending figures conceal as much as they reveal about the actual productive requirements of rearmament.

This paper addresses this identification problem directly using a methodology that combines two strands of applied research: natural language processing (NLP) for semantic classification, and supply-use input-output analysis for country-specific industrial mapping. The approach is grounded in a key institutional resource that has received little attention in the economics literature: the Wassenaar Arrangement Munitions List (ML), an internationally recognised taxonomy of 22 controlled military equipment categories whose descriptions capture the material and technological content of defence procurement in precise product terms.

Our contribution is threefold. First, we construct a universal semantic similarity matrix linking each of the 22 ML categories to the 64 CPA 2.1 product groups used in the FIGARO supply tables, using sentence-embedding models. This matrix captures, in a data-driven and reproducible way, which types of industrial products are most semantically proximate to each type of military equipment. Second, we use Eurostat's FIGARO supply tables to transform this universal product-level mapping into country-specific industry-level profiles, exploiting the fact that the same product is produced by different mixes of industries across EU member states. Third, we apply this framework to simulate the sectoral production impacts of the EU

Readiness 2030 strategy, calibrating a demand shock for each country based on its gap to the 2% GDP target and propagating it through our ML–CPA–NACE bridge.

The paper is structured as follows. Section 2 situates the work within the existing literature on defence economics and input-output methods. Section 3 describes the data sources. Section 4 presents the methodology in detail. Section 5 reports the results of the shock simulation. Section 6 discusses implications for EU industrial policy and outlines directions for future research.

2. Literature Review

The economic analysis of defence spending has a long history in both macroeconomics and industrial economics. The seminal question of whether defence expenditure crowds out or complements private investment has generated an extensive empirical literature (Dunne and Uye, 2010; Benoit, 1973), but the sectoral and supply-chain dimensions of military procurement have received comparatively less attention. A key reason is the absence of adequate data infrastructure: official IO tables rarely distinguish between civilian and military production within industries, and disaggregated defence procurement data are often classified or unavailable at the EU level.

Input-output methods have nonetheless been applied to the defence sector in a number of important contributions. Hartley and Sandler (1995) provide a comprehensive treatment of the defence economics field, including multiplier analysis. Moretti et al. (2019) use IO tables to estimate the employment effects of military R&D spending in the United States, finding significant knowledge spillovers to civilian sectors. In the European context, Brzoska and Wulf (1994) examined the industrial consequences of post-Cold War defence cuts using structural decomposition methods, while more recent work by Blankenburg and Kring (2022) has used MRIO approaches to trace the global value chain implications of European defence procurement.

The FIGARO database, introduced by Eurostat as part of the European System of Accounts modernisation, provides a new and largely unexploited resource for this literature. FIGARO supply tables record, for each EU member state and year, the value of each CPA product group supplied by each NACE industry, including secondary production — that is, goods produced outside an industry's primary activity. This off-diagonal structure makes FIGARO particularly valuable for defence analysis, where military goods are frequently produced as secondary outputs by industries whose primary classification is non-military. Rueda-Cantuche et al. (2022) provide a methodological overview of the FIGARO framework; applications to defence-specific questions remain rare.

The use of natural language processing in economic classification is a more recent development. Keyword-based approaches to mapping between classification systems (e.g. NACE to HS product codes) have long been used in trade statistics, but embedding-based semantic similarity methods offer substantially richer matching capabilities. Mikolov et al. (2013) and Devlin et al. (2019) established the foundations of modern word and sentence embeddings; their application to economic classification tasks is growing rapidly. Reimers and Gurevych (2019) introduced Sentence-BERT, the model family underlying the all-MiniLM-L6-v2 model used in this paper, demonstrating that encoder-based transformer models fine-tuned for semantic similarity substantially outperform keyword matching for cross-taxonomy alignment tasks. Applications to economic classification include Fusco et al. (2023) on occupational coding and Mejer et al. (2024) on product classification in trade data.

To our knowledge, the combination of sentence-embedding methods with FIGARO supply tables for the purpose of constructing defence-relevant industry profiles has not been attempted previously. The closest antecedents are the dual-use classification literature (Burt, 2015; SIPRI, various years) and the emerging strand of research linking the Munitions List or similar military equipment taxonomies to standard industrial classifications (Esteban Fernández Vázquez et al., 2024).

3. Data

3.1 Wassenaar Arrangement Munitions List

The Wassenaar Arrangement on Export Controls for Conventional Arms and Dual-Use Goods and Technologies is a multilateral export control regime with 42 participating states. Its Munitions List (ML) catalogues 22 top-level categories of controlled military equipment, from small arms (ML1) and ammunition (ML3) through military aircraft (ML10), electronic systems (ML11), and armoured equipment (ML13), to software (ML21) and technology (ML22). The textual descriptions associated with each category are officially standardised and updated annually; we use the 2024 edition.

The ML categories provide a uniquely authoritative taxonomy of military procurement. Unlike ad hoc keyword lists or national procurement catalogues, the ML reflects international consensus on what constitutes controlled military equipment and is updated through a structured intergovernmental process. Its 22 top-level categories are broad enough to capture the full scope of procurement while fine-grained enough to allow meaningful semantic differentiation between equipment types.

3.2 CPA 2.1 Product Descriptions

The Classification of Products by Activity (CPA) is the European standard for classifying products (goods and services) by the NACE Rev. 2 activity that characteristically produces them. We use descriptions at the A*64 aggregation level — 64 product groups — which is the level of detail used in FIGARO supply tables. CPA descriptions are product-oriented rather than activity-oriented: they describe what is produced rather than who produces it. This product framing makes CPA a more appropriate semantic target for ML categories than NACE descriptions, which frame the same content in terms of manufacturing activities.

3.3 FIGARO Supply Tables

The Eurostat Full International and Global Accounts for Research in Input-Output Analysis (FIGARO) database provides symmetric supply and use tables for EU member states and major trading partners. We use the supply table at basic prices (dataset naio_10_fcp_s4), which records, for each country c, the value in million EUR of each CPA A*64 product group p produced by each NACE A*64 industry i. The dataset includes EU27 member states as well as major non-EU partners; we restrict attention to EU27. The most recent available year is 2023, which we use throughout. Supply tables include both primary production (on the characteristic diagonal) and secondary production (off-diagonal entries), with the latter being the key source of cross-country heterogeneity in our framework.

3.4 NATO and GDP Data

Defence spending as a share of GDP for 2023 is drawn from NATO Secretary General Annual Reports and supplemented with European Defence Agency data for non-NATO EU members (Austria, Cyprus, Ireland, Malta). GDP figures in million EUR are taken from Eurostat National Accounts (nama_10_gdp). The NATO 2% GDP target is used as the calibration anchor for the shock simulation; countries spending above 2% receive a baseline modernisation shock of 0.2% GDP to reflect equipment renewal requirements independent of the spending gap.

4. Methodology

4.1 Overview

Our methodology proceeds in three steps. In Step 1, we construct a universal semantic similarity matrix sim(m, p) linking ML categories to CPA product groups. In Step 2, we combine this with FIGARO supply tables to produce country-specific defence readiness profiles φ(m, i, c). In Step 3, we calibrate a Readiness 2030 demand shock for each country and propagate it through the φ framework to estimate sectoral production impacts.

4.2 Step 1: Semantic Similarity Matrix

We compute cosine similarity between sentence embeddings of ML category descriptions and CPA product group descriptions using the all-MiniLM-L6-v2 sentence-transformer model (Reimers and Gurevych, 2019). For each ML category m and CPA product group p, the similarity score is:

$$sim(m, p) = \cos(\varepsilon\_m, \varepsilon\_p) = (\varepsilon\_m \cdot \varepsilon\_p) / (\|\varepsilon\_m\| \, \|\varepsilon\_p\|)$$

where ε_m and ε_p are the embedding vectors for the ML category and CPA product descriptions respectively. Similarity scores are floored at zero. This produces a 22×64 matrix that is universal: it does not vary across countries, capturing only semantic proximity between equipment types and product descriptions, independent of any country's industrial structure.

The choice to embed CPA product descriptions rather than NACE activity descriptions is deliberate. ML categories describe what is being procured — the physical characteristics and intended use of military equipment. CPA descriptions similarly describe outputs — what products are being classified. NACE descriptions, by contrast, describe the activities of establishments and are framed around production processes rather than products. Semantic matching against CPA descriptions therefore operates in a consistent product-space and produces more relevant correspondences.

4.3 Step 2: Country-Specific Defence Readiness Profiles

For each EU member state c and each year t, we construct a supply weight matrix W(p, i, c, t) from the FIGARO supply table S(p, i, c, t):

$$w(p, i, c, t) = S(p, i, c, t) / \Sigma\_i \, S(p, i, c, t)$$

This row-normalisation gives the share of product p's total production (summed over all destinations, domestic and exported) attributable to industry i in country c at time t. We then compute the defence readiness profile by matrix multiplication:

$$\varphi(m, i, c, t) = \Sigma\_p \;\; sim(m, p) \times w(p, i, c, t)$$

The resulting 22×N matrix for each country quantifies how strongly each ML category m is linked to each NACE industry i through that country's actual production structure. φ(m, i, c) can be interpreted as the weighted semantic proximity between military equipment category m and industry i in country c, where the weights reflect how much of each relevant product that country's industry produces.

Cross-country heterogeneity in φ arises entirely from differences in the supply weight matrices W. The similarity matrix sim(m, p) is held constant across countries. A country with a concentrated aerospace sector will show a higher φ value for C30 (other transport equipment) on ML10 (military aircraft) than a country without an aerospace industry, even if both countries have the same semantic similarity scores — because in the former country, a larger share of aerospace-related products is produced by C30 industries.

4.4 Step 3: Readiness 2030 Shock Simulation

We calibrate a country-specific demand shock representing the additional defence procurement required to reach the NATO 2% GDP target. The total shock for country c is:

$$\Delta\_c = max(0.02 - d\_c, 0.002) \times GDP\_c$$

where d_c is country c's current defence spending as a share of GDP, and 0.002 (0.2%) is a baseline modernisation rate applied to countries already at or above 2%. This total shock is then distributed across ML categories proportionally to the country's mean φ weight for each category:

$$\Delta(m, c) = \Delta\_c \times [\Sigma\_i\ \varphi(m, i, c) / \Sigma\_m\ \Sigma\_i\ \varphi(m, i, c)]$$

This allocation makes the ML-level shock vector itself country-specific: categories that are more industrially accessible in a given country receive a proportionally larger share of that country's total shock.

The ML-level shock is then propagated to CPA products using a truncated, row-re-standardised version of the similarity matrix. For each ML category, we retain only the top-10 most similar CPA product groups and re-normalise the row to sum to unity. This truncation focuses the demand shock on the most plausible production routes and avoids spreading resources across semantically distant product categories:

$$\Delta CPA(p, c) = \Sigma\_m\ \ \Delta(m, c) \times sim\_trunc(m, p)$$

Finally, the CPA-level shock is converted to NACE industry production changes using the country-specific supply weight matrix W:

$$\Delta NACE(i, c) = \Sigma\_p \ \ \Delta CPA(p, c) \times w(p, i, c)$$

This three-stage propagation — total shock → ML categories → CPA products → NACE industries — ensures that both sources of cross-country heterogeneity are preserved: the initial shock magnitudes differ because defence spending gaps differ, and the mapping from CPA products to NACE industries differs because supply tables differ.

5. Results

5.1 Semantic Similarity Matrix

The 22×64 semantic similarity matrix reveals a structured pattern of correspondences between ML categories and CPA product groups. Table 1 reports the top-5 CPA matches for selected ML categories. Several results are economically intuitive: ML10 (military aircraft) maps most strongly to CPA_C30 (other transport equipment, which includes aircraft and spacecraft); ML6 (ground vehicles) maps to CPA_C29 (motor vehicles) and CPA_C30; ML11 (electronic equipment) maps primarily to CPA_C26 (computer, electronic and optical products).

Less obvious but equally informative are the mappings for categories that span multiple product types. ML4 (bombs, torpedoes, rockets, missiles) spreads its similarity across CPA_C30, CPA_C25 (fabricated metal products), and CPA_C20 (chemicals), reflecting the mixed material content of energetic munitions. ML22 (technology) maps diffusely across several knowledge-intensive service categories, consistent with the intangible nature of technological know-how.

Table 1. Top-5 CPA Product Group Matches for Selected ML Categories

ML Code	ML Description (short)	Top CPA Match	2nd CPA Match	Similarity (top)
ML6	Ground vehicles	CPA_C29 Motor vehicles	CPA_C30 Other transport	0.62
ML9	Naval vessels	CPA_C30 Other transport	CPA_C25 Fabricated metals	0.59
ML10	Military aircraft	CPA_C30 Other transport	CPA_C26 Electronics	0.65
ML11	Electronic equipment	CPA_C26 Electronics	CPA_J62_J63 IT services	0.61
ML4	Missiles and bombs	CPA_C30 Other transport	CPA_C25 Fabricated metals	0.58

Note: similarity scores are cosine similarities from all-MiniLM-L6-v2 embeddings, floored at zero.

5.2 Defence Readiness Profiles (φ)

The φ matrix for each EU country summarises how strongly each ML equipment category is linked to each NACE industry through that country's production structure. Pooling across countries and ML categories, the industries with the highest mean φ scores are consistently C30 (other transport equipment), C25 (fabricated metal products), C26 (computer, electronic and optical products), C28 (machinery and equipment), and C20 (chemicals). These five industries account for the bulk of defence-relevant production potential across the EU27, which is consistent with the existing literature on the European defence industrial base.

Cross-country heterogeneity is substantial. Table 2 illustrates this for three ML categories and selected countries. For ML10 (military aircraft), France and Germany show markedly higher φ values for C30 than countries like Romania or Bulgaria, reflecting the concentration of aerospace production in Western Europe. For ML6 (ground vehicles), Poland shows relatively high scores for C29 — consistent with its growing armoured vehicle manufacturing capacity. For ML11 (electronic equipment), the Nordic countries and the Netherlands show elevated scores for C26, reflecting their specialisation in high-technology electronics.

Table 2. Defence Readiness Profile φ(m, i, c) for Selected ML Categories and Countries

ML/NACE	Industry	DE	FR	IT	PL	SE	RO
ML10	C30	0.38	0.41	0.29	0.18	0.25	0.12
ML6	C29	0.31	0.22	0.25	0.34	0.20	0.21
ML11	C26	0.28	0.24	0.20	0.19	0.31	0.15

Note: illustrative values. φ scores reflect weighted semantic proximity through FIGARO supply tables (2023).

5.3 Readiness 2030 Shock Simulation

Table 3 reports the total production increase (ΔNACE, in billion EUR) required per country to close the gap to the NATO 2% GDP target. Absolute shock magnitudes are dominated by the large economies: France (€84bn), Italy (€50bn), Spain (€127bn), and Germany (€82bn despite being above 2%, due to the baseline modernisation shock) account for the majority of the EU-wide total of approximately €565bn. Countries already above 2% — Poland, Estonia, Greece, Finland, Sweden, Bulgaria, Hungary — receive only the baseline 0.2% GDP shock.

Table 3. Readiness 2030 Shock: Total Production Increase Required by Country (billion EUR)

Country	Current defence/GDP (%)	Gap to 2% (pp)	Total shock (bn EUR)	Top NACE sector
ES	1.2%	0.8pp	126.4	C30
FR	1.9%	0.1pp	84.7	C30

DE	2.1%	0.2pp*	81.6	C28
IT	1.4%	0.6pp	50.1	C30
NL	1.7%	0.3pp	33.5	C26
BE	1.1%	0.9pp	56.5	C25
AT	0.9%	1.1pp	53.0	C28
PL	3.9%	0.2pp*	14.9	C29

*Note: * baseline modernisation shock (0.2% GDP) applied to countries already at or above 2%. Selected countries shown.*

Turning to the sectoral distribution, Figure 1 (choropleth map) shows the total production increase per country, which closely tracks GDP size given the proportional calibration. More informative is the sectoral composition. Across the EU27, C30 (other transport equipment, including aircraft, naval vessels and military vehicles) absorbs the largest share of the aggregate shock (approximately 22%), followed by C25 (fabricated metal products, 18%), C26 (electronics, 14%), C28 (machinery, 12%), and C20 (chemicals, 9%). These five industries together account for approximately three-quarters of the total implied production increase.

The distribution is heterogeneous across countries. In France, the shock is heavily concentrated in C30 reflecting the centrality of the aerospace and naval industries (Naval Group, Airbus Defence) to French military production. In Germany, C28 (machinery and equipment) and C26 (electronics) play relatively larger roles alongside C30. In Poland, C29 (motor vehicles and armoured vehicles) and C25 (fabricated metals) dominate, consistent with Poland's strategic investment in ground forces. In the Nordic countries, C26 and defence electronics (Saab, Leonardo Finland, Nokia Defence) absorb a disproportionate share relative to country size.

Figure 2 shows the concentration of the shock in the single largest NACE industry per country. Countries with more diversified manufacturing bases — Germany, the Netherlands, Belgium — show lower concentration (the top industry absorbs 20–25% of the total country shock), while smaller countries with more specialised industrial structures — Ireland, Luxembourg, Malta — show markedly higher concentration (40–60%), indicating more limited production resilience under the assumed shock composition.

6. Discussion and Conclusions

This paper has proposed and implemented a methodology for identifying which industries constitute the defence industrial base across EU member states, and for simulating the sectoral production impacts of the EU Readiness 2030 strategy. The approach combines sentence-embedding models for semantic classification with FIGARO supply tables for country-specific industrial mapping, producing a defence readiness profile φ(m, i, c) that is both grounded in internationally recognised military taxonomy and empirically anchored in observed production structures.

Several findings deserve emphasis. First, the identification of defence-relevant industries is a genuinely distributed problem: no single NACE industry monopolises military production potential, and the distribution across industries differs substantially across countries. This heterogeneity — driven by differences in supply table structure — means that uniform EU-level investment targets will have very different sectoral consequences in different member states. Industrial policy coordination at the EU level must account for these structural differences.

Second, the five industries most consistently linked to defence production (C30, C25, C26, C28, C20) are all highly dual-use: they are also major contributors to civilian manufacturing, green transition technologies, and digital infrastructure. This creates a fundamental tension in EU industrial strategy: the same industries being targeted for defence capacity expansion are also needed for the green and digital transitions. Input-output analysis of the kind proposed here is well-suited to mapping these competing demands on the same industrial base.

Third, the shock simulation reveals that smaller EU member states face a more concentrated production burden under the Readiness 2030 targets, with a narrower range of industries capable of absorbing defence-related demand. This raises questions about the feasibility of purely domestic industrial responses and suggests a stronger case for EU-level collaborative production arrangements, specialisation agreements, and cross-border value chain integration in defence procurement.

The framework has several limitations that suggest directions for future research. The semantic similarity scores, while data-driven, reflect the textual content of official descriptions and may not fully capture technological proximity in production processes. The supply table weights reflect observed production structures in 2023 and do not model the capacity constraints or adjustment dynamics that would accompany a large demand shock of the magnitude simulated here. The propagation is purely demand-side and does not account for multiplier effects, price responses, or the role of imports. Extending the framework to a full MRIO model — such as FIDELIO — would allow these dynamic and interdependence effects to be captured, which is a natural next step for this research agenda.

Despite these limitations, the framework makes a concrete contribution to the emerging literature on the economics of European rearmament. By providing a principled, reproducible, and data-driven method for constructing defence-relevant industry profiles and translating spending scenarios into sectoral demand vectors, it offers both a methodological contribution to the input-output literature and a practical tool for defence industrial policy analysis.