

National and Subnational Social Accounting Matrices for the European Union and the United Kingdom, 2017–2022: A Harmonised Dataset Based on Official Accounts with an Annual Update Framework

Zein Khraizat¹, Olivier Dessens¹, Alvaro Calzadilla¹

¹ Institute for Sustainable Resources, University College London,
14 Upper Woburn Place, London WC1H 0NN, United Kingdom

* Corresponding author: Zein Khraizat (z.m.khraizat@ucl.ac.uk)

Keywords: *Social Accounting Matrices, NUTS2, Regional SAMs, Interregional Linkages, EU27, United Kingdom, Bilateral Trade, Agriculture focused SAMs*

Abstract

Economy wide assessments of climate change, trade linkages, and structural transformation require internally consistent representations of production, income generation, and trade flows. Social Accounting Matrices (SAMs) provide such a framework, yet consistent and regularly updated SAMs at both national and subnational levels remain limited despite advances in international input output databases such as Eurostat's FIGARO and national sources including the UK Office for National Statistics (ONS). To address this gap, we construct a harmonised dataset of national and subnational SAMs for the European Union (EU27) and the United Kingdom covering the period 2017 to 2022, fully based on official data and benchmarked to FIGARO and ONS accounts. At the national level, the dataset provides country SAMs linked through bilateral trade flows with 17 major EU trading partners and a Rest of the World account, while retaining detailed sectoral structure and extending agricultural representation using the Economic Accounts for Agriculture. Building on this benchmark, the database is regionalised to the NUTS2 level, covering 244 EU regions and 41 UK regions within a fully consistent national and global accounting framework. Regional production, value added, and demand are allocated using official regional statistics, while interregional and international trade flows are estimated using a gravity-based approach constrained to national totals. The framework maintains consistency across production, trade, and institutional accounts, including explicit treatment of trade margins and taxes in line with GTAP conventions. The resulting dataset provides a consistent empirical foundation for computable general equilibrium modelling of climate impacts, agricultural productivity shocks, and trade spillovers, and represents ongoing work to be further refined and updated through an annual release framework.

1. Introduction

The economic implications of structural change, trade integration, and environmental pressures are transmitted through complex networks of production, income generation, and exchange that link sectors, regions, and institutions. Analysing these interdependencies requires data frameworks capable of capturing the circular flow of income and expenditure across the entire economy, while preserving the structural relationships that govern how activity in one part of the system affects others [1, 2]. In practice, however, empirical representations of the economy remain fragmented across scales: global and inter-country input-output databases provide internally consistent national accounts and trade linkages, while regional statistics capture only partial information on production and income, without the full accounting structure required for economy-wide analysis. This disconnect limits the capacity to assess regionally differentiated impacts, particularly in sectors such as agriculture where production structures, resource endowments, and exposure to environmental change vary significantly across space [3].

SAMs provide a comprehensive framework to address this limitation by extending input-output structures to incorporate factor incomes and institutional accounts within a unified representation of the circular flow of income [1, 4, 5]. By reconciling production, consumption, and distributional relationships, SAMs enable consistent economy-wide analysis and form the empirical backbone of computable general equilibrium models [2, 5]. However, despite their analytical advantages, the availability of harmonised SAMs remains constrained, particularly at the subnational level [3, 6]. While recent advances in international databases, such as those developed by Eurostat through the FIGARO framework, have improved the consistency and timeliness of national accounts and inter-country trade linkages, these datasets are not designed to capture within-country spatial heterogeneity [7–9].

Extending nationally consistent accounting systems to the regional scale is fundamentally a problem of statistical reconciliation across multiple dimensions [7, 8, 10, 11]. National totals for output, value added, intermediate use, final demand, and trade must be distributed across heterogeneous territories in a manner that preserves accounting identities, respects observed regional structures, and remains consistent with the technological relationships embedded in the national benchmark tables. However, regional allocation alone is insufficient for many analytical applications, because the required level of sectoral detail is often suppressed in aggregate national accounts or unevenly reported in regional statistics [12, 13]. The challenge is therefore not only spatial, but also sectoral: policy analysis frequently depends on distinctions between activities that are concealed within broad accounting aggregates [14]. This is especially relevant for agriculture, where production technologies, land use, climatic conditions, and market orientation differ markedly across regions, generating forms of heterogeneity that cannot be represented adequately through nationally averaged structures. Constructing credible regional accounts therefore requires an integrated framework capable of combining macro-consistent benchmarks with auxiliary regional information while retaining economically meaningful sectoral resolution [15].

This paper addresses these challenges by developing a reproducible methodology for constructing harmonised Social Accounting Matrices at both national and subnational scales for the EU27/UK. The approach is fully anchored in official statistics, combining inter-country supply and use tables with regional accounts and

sector-specific data to generate a consistent multi-scale database. Starting from nationally balanced accounts, production, value added, and demand are allocated to the NUTS2 level using official regional indicators, while maintaining strict consistency with national aggregates. Within this framework, particular emphasis is placed on the explicit representation of agriculture at the regional level, allowing the resulting SAMs to capture spatial variation in agricultural production and its interaction with the broader economic system.

The resulting dataset provides a coherent and extensible empirical foundation for economy-wide analysis at both national and regional scales. By integrating detailed sectoral structure with consistent accounting across spatial levels, the framework enables the analysis of regionally differentiated dynamics within a unified system, and establishes a basis for future extensions and regular updates aligned with official statistical releases.

2. Methodology

The dataset was constructed to extend nationally consistent macroeconomic accounts to the subnational scale while preserving the accounting coherence of the benchmark system. The objective was to generate a harmonised set of national and regional Social Accounting Matrices (SAMs) for the EU27 and the United Kingdom that combine the internal consistency of official supply and use statistics with the spatial resolution required for regional analysis. The workflow follows a top-down regionalisation strategy in which nationally balanced accounts are treated as the reference structure and additional regional and sectoral detail are introduced through controlled statistical disaggregation.

The construction procedure was implemented in three sequential stages. First, national benchmark accounts were regionalised to NUTS2 territories using official regional indicators. Second, regional production, demand, and factor income structures were refined using sector-specific datasets and constrained updating procedures that preserve consistency with the benchmark accounts. Third, all estimated flows were assembled into an interconnected multi-regional SAM system and reconciled through balancing procedures to ensure exact consistency across regions, sectors, institutions, and national totals.

All transformations were implemented as constrained adjustments of the benchmark accounts rather than as independent estimations. This ensures that the final database reproduces official macroeconomic totals while incorporating regional heterogeneity and additional sectoral detail. The resulting framework combines the statistical robustness of official accounts with the analytical resolution required for regional policy analysis, impact assessment, and economy-wide modelling.

2.1 Data architecture

FIGARO provides the benchmark accounting structure for all countries represented in the dataset. The system covers the 27 Member States of the European Union, the United Kingdom, the United States, and 17 major trading partners of the EU, thereby embedding domestic economies within a wider international production and trade network. Within this harmonised inter-country framework, FIGARO defines the official control totals for output, intermediate consumption, final demand, value added, and bilateral trade flows [9]. This broad geographic coverage is particularly important because regional economies are not closed systems, but are shaped by national and international supply-chain linkages, external demand, and import

dependence [13]. Compiled from national accounts, supply and use tables, and international trade statistics under common European statistical standards, FIGARO therefore provides a coherent and internationally comparable macroeconomic benchmark for all subsequent stages of estimation and regionalisation.

Regionalisation of national aggregates was based primarily on NUTS2 gross value added (GVA) statistics, which provide the principal indicator for the spatial distribution of sectoral economic activity [16]. Agricultural output, subsidies, and factor income components were informed by the Economic Accounts for Agriculture (EAA) [17]. Farm-level production technologies and expenditure patterns were informed by the Farm Accountancy Data Network (FADN) [18]. Commodity-specific demand structures, particularly for agricultural uses and final demand composition, were supported by CAPRI data [19]. Spatial information on irrigated and rainfed production systems was derived from the IFPRI SPAM database.

2.2 Construction procedure

The database was generated through three sequential components that translate benchmark national accounts into a balanced multi-regional SAM system.

Regionalisation of national accounts. National aggregates from FIGARO were first allocated to NUTS2 regions using official regional indicators. Sector-specific gross value added was used as the principal allocation key for output and value added, as it provides the most consistent observable measure of the spatial distribution of productive activity where complete regional supply and use tables are unavailable. For each country and sector, regional estimates were constrained to sum exactly to the corresponding national totals. Where regional indicators were incomplete or unavailable, supplementary information from adjacent years, related sectors, or statistically comparable observations was used to construct provisional values, which were subsequently rescaled to preserve benchmark totals. This stage produced a complete regional distribution of production and income accounts consistent with the national framework.

Refinement of regional production and demand structures. The initial regional accounts were then enriched using sector-specific datasets to introduce detail not directly available in the benchmark tables. The aggregate agricultural sector was disaggregated into detailed crop and livestock activities using production shares from the Economic Accounts for Agriculture [17]. Benchmark input structures inherited from the national supply and use tables were updated through constrained entropy-based procedures to reflect regional differences in production technologies, input intensity, and land use while remaining close to the benchmark coefficients. Regional final demand was refined by allocating expenditure categories across the expanded commodity structure using CAPRI demand information, particularly for agricultural commodities. Factor income accounts were further specified through the separate treatment of land rents, allocated across agricultural activities using observed land use and production patterns. Agricultural output was also differentiated into irrigated and rainfed systems using spatial irrigation shares from the SPAM database, enabling the derivation of associated water use accounts [20].

Compilation, balancing and SAM construction. All estimated flows were subsequently assembled into an interconnected multi-regional SAM system in which each regional economy records transactions between production activities, commodities, factors of production, households, government, capital accounts, and the rest of the world. Regional accounts were linked through domestic and international trade flows, allowing the full system to capture interregional dependence within a unified accounting framework. Preliminary estimates were then reconciled using constrained optimisation procedures based on cross-entropy principles, which adjust cell values minimally relative to their prior estimates while enforcing all accounting identities simultaneously. These constraints included equality between row and column totals within each regional SAM, consistency between the sum of regional accounts and national benchmark totals, and coherence of trade flows across origin and destination accounts. The final dataset is therefore internally balanced and exactly consistent with the benchmark macroeconomic accounts.

3. Conclusion

This work develops a transparent and reproducible framework for constructing harmonised national and NUTS2-level Social Accounting Matrices that extend official macroeconomic accounts to the regional scale while preserving full accounting consistency. The resulting database combines the benchmark coherence of FIGARO with the spatial resolution required to analyse heterogeneous regional economies within an integrated European setting.

A central contribution of the framework is the enhanced representation of agriculture. By combining regional allocation procedures with coefficient updating informed by EAA, FADN, and CAPRI data, the database captures differences in production structures, commodity composition, and factor use across regions and agricultural activities. The explicit treatment of land rents, together with the distinction between rainfed and irrigated systems and their associated water accounts, further broadens the analytical scope of the dataset by linking economic activity to resource use and environmental constraints.

Methodologically, the study demonstrates how multiple statistical sources can be reconciled within a single top-down accounting framework through constrained disaggregation and balancing procedures. The resulting dataset extends nationally consistent accounts to the regional scale where complete subnational tables are unavailable, while maintaining exact consistency with published macroeconomic totals. It provides an empirical basis for regional policy assessment, trade and supply-chain analysis, structural change, and computable general equilibrium modelling, and, because the workflow is modular and replicable, can be updated with new statistical releases, expanded with additional environmental satellite accounts, and adapted to further sectoral refinement.

References

- [1] Jeffrey Round. Constructing sams for development policy analysis: lessons learned and challenges ahead. *Economic Systems Research*, 15(2):161–183, 2003.
- [2] Ronald E Miller and Peter D Blair. *Input-output analysis: foundations and extensions*. Cambridge university press, 2009.

- [3] Lucía Bolea, Rosa Duarte, Geoffrey JD Hewings, Sofía Jiménez, and Julio Sánchez-Chóliz. The role of regions in global value chains: an analysis for the european union. *Papers in Regional Science*, 101(4): 771–795, 2022.
- [4] Graham Pyatt and Jeffrey I Round. Social accounting matrices for development planning 1. *Review of Income and Wealth*, 23(4):339–364, 1977.
- [5] Clemens Breisinger, Marc Thomas, and James Thurlow. *Social Accounting Matrices and Multiplier Analysis: An Introduction with Exercises*. Food Security in Practice Technical Guide 5. International Food Policy Research Institute, Washington, DC, 2009.
- [6] Bjarne Madsen and Chris Jensen-Butler. Spatial accounting methods and the construction of spatial social accounting matrices. *Economic Systems Research*, 17(2):187–210, 2005. doi: 10.1080/09535310500114994.
- [7] Jose Manuel Rueda-Cantuche, Jorge Lopez Alvarez, and Luis Antonio Galiano Bastarrica. FIGARO-REG 2017 10 sectors. European Commission, Joint Research Centre [Dataset], 2017. URL <http://data.europa.eu/89h/df29c8d-b85b-41fa-9cb7-7289c7324937>. PID: <http://data.europa.eu/89h/df29c8d-b85b-41fa-9cb7-7289c7324937>.
- [8] Siyu Huang and Pantelis Koutroumpis. European multi regional input output data for 2008–2018. *Scientific data*, 10(1):218, 2023.
- [9] Eurostat. Figaro database—eu inter-country supply, use and input-output tables. Database, 2025. URL <https://ec.europa.eu/eurostat/web/esa-supply-use-input-tables/database#figarotables>. Accessed 2026.
- [10] Manfred Lenzen, Daniel Moran, Keiichiro Kanemoto, and Arne Geschke. Building eora: a global multi-region input–output database at high country and sector resolution. *Economic systems research*, 25(1):20–49, 2013.
- [11] Eduardo Rodrigues Sanguinet, Adelar Fochezatto, and Cristian Gonzalez Santander. Regionalization of input–output matrices with limited information: Application to the state of rio grande do sul, brazil. *Regional Science and Environmental Economics*, 2(4):31, 2025.
- [12] N Szabó. Methods for regionalizing input-output tables regional statistics 5 (1): 44–65, 2015.
- [13] Anthony T Flegg and Timo Tohmö. The regionalization of national input–output tables: a review of the performance of two key non-survey methods. *Applications of the input-output framework*, pages 347–386, 2018.
- [14] Barbara Hutniczak. Efficient updating of regional supply and use tables with the national-level statistics. *Journal of Economic Structures*, 11(1):16, 2022.
- [15] Marcel Adenäuer, Wolfgang Britz, and Andrea Zimmermann. Spatial heterogeneity of the agricultural sector in economic models, 2016.

- [16] Eurostat. Gross value added by nuts 2 region. <https://ec.europa.eu/eurostat/databrowser/view/tgs00003/default/table?lang=en>, 2026. Source: Eurostat dataset tgs00003, accessed 14 April 2026.
- [17] Eurostat. Economic accounts for agriculture by nuts 2 region - values at current prices. https://ec.europa.eu/eurostat/databrowser/product/view/aact_eaa01_r?lang=en, 2026. Source: Eurostat dataset https://doi.org/10.2908/ACT_EAA01_R, accessed 14 April 2026.
- [18] European Commission. Farm accountancy data network (fadn) public database. <https://agridata.ec.europa.eu/extensions/FarmEconomyFocus/FADNDatabase.html>, 2024. Accessed April 14, 2026.
- [19] R. Koeble, A. Leip, M. Kempen, et al. Eu wide disaggregated capri model data: Crops, livestock, nitrogen in- and outputs (timeseries 2000-2018). *Scientific Data*, 2026. doi: 10.1038/s41597-026-06919-8. URL <https://doi.org/10.1038/s41597-026-06919-8>.
- [20] International Food Policy Research Institute (IFPRI). Global Spatially-Disaggregated Crop Production Statistics Data for 2020 Version 2.0, 2024. URL <https://doi.org/10.7910/DVN/SWPENT>.