Interregional Input-Output Table for Brazil: A Data-Driven Estimation Using the Brazilian Electronic Invoice System

Topic: Regional Input-output Modelling (2)
Author: Joao Maria de Oliveira
Co-Authors: Patieene Alves Passoni

This study aims to estimate an Interregional Input-Output Table (IIOT) for Brazil using a data-driven methodology based on the Brazilian Electronic Invoice System for 2018. The increasing availability of administrative data, such as fiscal records and electronic transactions, enables a more refined and precise approach to estimating these tables, addressing historical limitations related to the scarcity and outdated nature of traditional data sources. Improving input-output table estimation methodologies significantly contributes to the development of more robust economic analyses and evidence-based policymaking.
Originally designed for national applications, Input-Output Models have been applied to subnational geographical units since the second half of the last century. Two specific characteristics of the regional dimension make the distinction between national and regional input-output models essential: the production technology of each region is unique and may be either like or significantly different from that recorded in the national input-output table; and smaller economies are more dependent on external markets, making imported and exported demand and supply components increasingly relevant. A key advantage of interregional models is their ability to account for spillover effects, which allow for the assessment of how an increase in demand in one region affects others while also considering feedback effects on the region where the initial increase occurred.
This research aligns with the recent trend of producing estimates based on official administrative records, offering a more accurate representation of economic reality. It estimates Brazilâ€™s interregional input-output table using data from the Brazilian Electronic Invoice System (NF-e). NF-e data captures most business-to-business transactions, particularly those involving interregional trade, addressing a significant data gap in interregional input-output table estimation. The table was estimated at the States level, ensuring compatibility with the System of National Accounts (SNA) and the System of Regional Accounts (SRA) of the Brazilian Institute of Geography and Statistics (IBGE).
This estimation strategy enables the exploration of key research questions, such as how NF-e data can be used to construct a detailed and accurate interregional input-output table, what is the interregional structure of this table as revealed by NF-e data and how it differs from previous estimates, and what is the input-output table of each of Brazilâ€™s 27 states and how it differs from isolated state-level estimations. The methodology involves data cleaning and preprocessing from an extraction of the Brazilian NF-e database for 2018. The study employs product codes from the Mercosur Common Nomenclature (NCM), which aligns with the International Harmonized System (HS), along with seller and buyer sector codes and tax codes that classify transactions. A critical step is applying the Broad Economic Categories (BEC) classification to determine whether products should be classified as intermediate consumption, final consumption (households/government), or gross fixed capital formation.
The key variables include NCM/HS product codes, seller and buyer sector codes, legal classification (public or private), transaction values, product tax information, fiscal classification of the transaction, and geographical data (state of origin and destination). The resulting table is designed for seamless integration with both the SNA and SRA.
As a result, the study presents an IIOT covering 27 states and 68 sectors, capturing also the flow of exports and imports between regions. A version was also estimated considering the 5 Brazilian macro-regions. The novelty of this research lies in the pioneering use of comprehensive and detailed NF-e data to estimate intersectoral and interregional transactions and their economic interactions. Unlike previous interregional input-output table estimations for Brazil, which often relied

on national data and approximations, this study uses NF-e data to capture actual interregional trade flows, providing a more accurate representation of economic reality.
The study also discusses the challenges and limitations associated with using administrative data, including data quality issues and classification ambiguities. However, the approach offers a significant advantage by utilizing a data-driven methodology that leverages the detailed and extensive coverage of NF-e data. The research contributes to the advancement of methodologies for regional economic analysis and intersectoral linkages estimation, providing valuable insights for evidence-based policymaking. Potential future research directions include expanding the time series analysis, integrating other data sources for enhanced accuracy, and exploring the dynamic interactions between regions.

Key words: Input-Output Analysis, Electronic Invoices, Brazil, Regional Analysis, Big Data.