Building an alliance to map global supply networks

A better understanding is key to resolving major societal challenges

By **Anton Pichler[1,2], Christian Diem[1,2], Alexandra Brintrup[3], François Lafond[4], Glenn Magerman[5], Gert Buiten[6], Thomas Y. Choi[7], Vasco M. Carvalho[8,9,10], J. Doyne Farmer[4,11], Stefan Thurner[1,12,11,13]**

The global economy consists of more than 300 million firms, connected through an estimated 13 billion supply links (see supplementary materials), that produce most goods and services. It has long been unthinkable to analyze the world economy at the firm level, even less so its intricate network of supply chain linkages. This blind spot has left us ill-prepared to make fast and well-informed decisions, begetting, for example, prolonged shortages in raw materials and critical medical supplies during the COVID-19 pandemic. Now, the availability of new data and recent methodological advances allow us to reconstruct a large share of the global firm-level supply network. Since mapping this network is likely to continue to improve, it is essential to initiate a discussion about responsible management and effective use of this data for the global public good. This requires new collaborative efforts between nations, their public institutions, international organizations, the private sector, and scientists.

Potential applications of a global supply chain map include monitoring and improved management of the green transition, reducing tax evasion and corruption, strengthening human rights through supply chain transparency, identifying and monitoring systemic risks and systemically important firms, and the design of globally secure basic provisioning systems for food and medication. Toward such ends, research has contributed to a better understanding of the functioning of supply chains, yet has been constrained by data limitations. While highly granular data are available for single "focal" firms where the direct suppliers and customers are known, such data are not connected to the rest of the economy, and no network perspective is possible. In contrast, when looking at how supply chains affect economy-wide phenomena such as GDP, business cycles, or inflation, one is typically constrained to highly aggregated data in the form of input-output relationships between a few dozen industrial sectors.

Until now, such data limitations made it impossible to integrate granular product- and firm-level expertise with macro-level, economy-wide perspectives. Recently, high-resolution maps of firm-level supply networks have been charted for individual economies (1). However, since they are not connected, the global image – the one that really counts in a globalized economy – remains fragmented. Even national maps are typically unavailable to researchers and policymakers, limiting their potential usefulness for addressing societal problems.

A NEED FOR GLOBAL NETWORK DATA

In 2021, supply disruptions caused an estimated 2% loss of global GDP (roughly 1.9 trillion USD) and substantially contributed to high inflation (*2*). In a globalized economy, local shocks to individual firms can spread to geographically distant firms through several tiers in the supply chain. Spotting such spreading dynamics ahead of time is impossible with aggregate, industry-level data or granular but incomplete supply chain data.

The potential for modeling the propagation of economic shocks in a detailed manner has been shown for individual countries for which large-scale firm-level supply network data is available. For example, in 2011, the Great East Japan Earthquake affected only four prefectures directly, but the economic shocks propagated throughout the economy. Based on detailed Japanese supply network data of several million firm-level supply links, it was possible to model the indirect country-wide economic impacts in much detail (*3*).

A new generation of economic models becomes possible when calibrated to granular supply network data, enabling us to address "what-if" scenarios for direct use in policy interventions at the micro- and macro-level. For example, using the detailed Japanese dataset, researchers predicted a recession of 2.2% with a probability exceeding 10% if a Tōkai Earthquake with magnitude 8+ occurs – which is considered likely in the coming decades (*3*).

In some countries, value-added tax (VAT) data is collected on the level of individual trades, allowing us to recover virtually all domestic business-to-business trades. This data has been used to quantify the economic importance of individual firms by modeling their direct and indirect macroeconomic impact in case of failure (*4*). Thus, analogously to financial networks, we could leverage detailed supply network data to monitor the systemic risk levels posed by individual firms and incorporate this information into economic regulation.

Complex supply relationships are also highly susceptible to tax fraud. Between 2016 and 2019, the annual VAT-related tax gap has been estimated to EUR 120 billion within the European Union (EU). "Carousel fraud" alone, which involves the intentional circulation of goods among companies and countries to avoid VAT payments, accounts for around

[1]Complexity Science Hub, Vienna, Austria. [2]Vienna University of Economics and Business, Vienna, Austria. [3]Institute for Manufacturing, University of Cambridge, Cambridge, UK. [4]Institute for New Economic Thinking and Mathematical Institute, University of Oxford, Oxford, UK. [5]European Center for Advanced Research in Economics and Statistics, Université Libre de Bruxelles, Brussels, Belgium. [6]Centraal Bureau voor de Statistiek, Netherlands. [7]Department of Supply Chain Management, W. P. Carey School of Business, Arizona State University, Tempe, USA. [8]Faculty of Economics, University of Cambridge, UK. [9]Centre for Economic Policy Research, London, UK. [10]Alan Turing Institute, London, UK. [11]Santa Fe Institute, Santa Fe, USA.[12]Section for Complex Systems, Medical University of Vienna, Vienna, Austria. [13]Supply Chain Intelligence Institute Austria, Vienna, Austria. Email: pichler@csh.ac.at

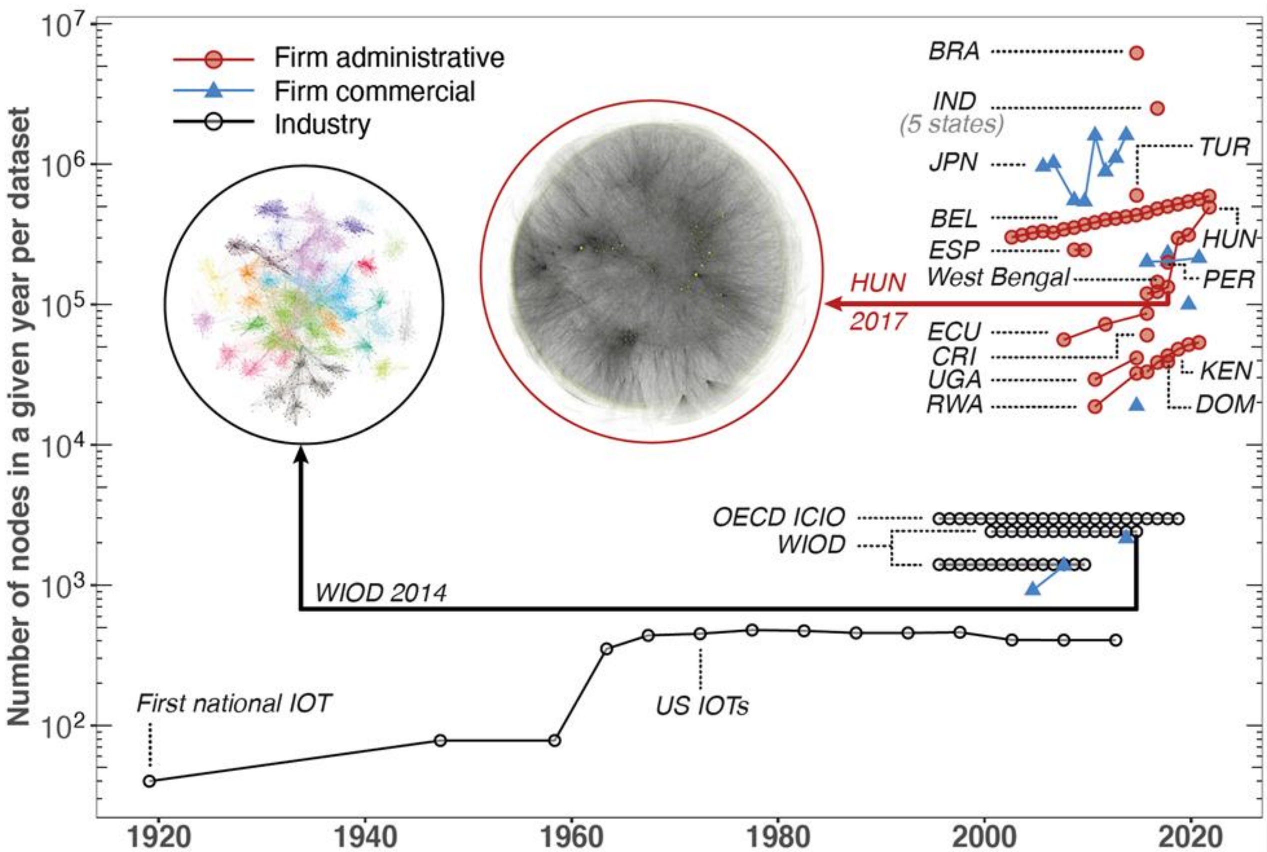

Figure 1: The rise of large-scale firm-level supply network data. The figure shows an explosion in the number of nodes in supply network datasets in recent years. In firm-level datasets, nodes represent firms, while in input-output data, nodes represent economic sectors or sector-country pairs. Data points are connected through a line if the dataset is observed over multiple years. We labeled administrative datasets with their ISO 3-letter country code. Administrative datasets shown are derived from VAT, except for Brazil, which has been collected from payment data over multiple years. For national industry-level data, the US input-output tables (IOTs) are shown, which are among the most detailed in the world (ca. 400 sectors). WIOD and OECD ICIO represent examples of inter-country IOTs primarily derived bottom-up from national IOTs and trade data. TSR stands for Tokyo Shoko Research.

EUR 50 billion in lost tax revenues annually. The ability to measure individual supply linkages between firms could be immensely valuable in detecting and preventing such fraudulent activities.

As policymakers increasingly recognize, supply chains are critical for ensuring compliance with tax laws, human rights, and environmental standards. Recent policy initiatives such as the U.S. "Uyghur Forced Labor Prevention Act" or the EU "Supply Chain Act" aim to hold companies accountable for monitoring and upholding social and environmental standards throughout their supply chain.

Managing the transition to a carbon-neutral economy requires a detailed understanding of supply chain relationships. Due to the lack of data, it is currently impossible to accurately assess and monitor the carbon intensity of products and indirect emissions caused by individual firms. Deploying new green technologies on a large (global) scale while phasing out fossil technologies will drastically "rewire" the supply network, resulting in heterogeneous geographical impacts on employment, energy prices, and financial stability. These changes will create new geopolitical dependencies and shift vulnerabilities in economic production by introducing new critical materials.

A NEW ERA OF MICRO-DATA

Until very recently, supply network data has been limited to a few hundred firms or industries (traditional input-output tables). Only in the past decade, a supply network data revolution has started with several independent lines of progress, including payment data, VAT data, and various reconstruction methods. Through these efforts, the scope of supply chain data could be increased by several orders of magnitude (see Fig. 1).

The ongoing expansion of supply chain data is also driven by the private sector. As demonstrated by recent investments in supply chain analytics firms worth hundreds of millions USD, a better understanding of supply dependencies is highly valued information. Datasets collected by firms providing business intelligence services can be extensive, including over 200K firms with more than 500K supply linkages (*1*), and have proven vastly useful for reconstructing supply chain relationships. These datasets are derived from a wide range of sources, including credit rating agencies, firm disclosures on public filings, business reports, capital market presentations, and press releases. Limitations of commercial datasets include a bias towards large publicly listed firms, uncertain quality assessment, and the lack of methodological transparency (*5*).

To overcome some of these issues, payment data and bank transaction data have recently been tapped to reconstruct supply linkages between firms (*6*). However, this data is typically not readily available for research or supporting policy making, and it can be challenging to reliably extract supply dependencies from payment data.

The most complete information on firm-level supply networks that exists today comes from administrative VAT records. Out of the 170 countries that collect VAT, researchers have mapped virtually all domestic trades between firms for Belgium, Chile, Kenya, Turkey, Ecuador, Costa Rica, Uganda, Hungary, Spain, Rwanda, multiple states of India, and the Dominican Republic, and similar efforts are ongoing for further countries. These datasets represent an unprecedented window into the microstructure of buyer-customer relationships that enable us for the first time to characterize entire national economies as complex supply networks (*7*, *8*).

These datasets show that individual firms can have tens of thousands of suppliers, which themselves are connected to a vast network of supply linkages. Because firms can only observe their direct suppliers and customers, it is very difficult for them to reconstruct the upper tiers of their supply chains. Instead, utilizing already collected tax data could provide a very extensive picture without additional administrative burdens for companies.

A MAPPING BLUEPRINT

Even though VAT data enables us to reconstruct complete domestic supply networks at the firm level, it is necessary for many, if not most, essential applications to combine these for different countries. Securing the provision of critical goods, monitoring human rights, fighting cross-border tax fraud, and monitoring the carbon footprint of individual products will remain ineffective without a granular, comprehensive, and international map of supply linkages.

Extending the detailed domestic supply networks across country borders requires regionally and internationally coordinated strategies at the same time. First, it is necessary to collect the domestic firm-to-firm trades via granular VAT records, e.g., using e-invoicing. Such systems can even reduce administrative burdens borne by firms and substantially contribute to tax compliance (*9*). Thus, treasuries and tax offices should have strong incentives to do this. Several countries have already successfully implemented efficient reporting standards that can serve as role models for others (see SM). Notably, among those countries, there are several small and developing economies with limited resources, demonstrating the technical and economic feasibility of collecting firm-level supply network data through VAT records.

The next step requires connecting different countries based on trade data. In the EU, the firm-level trade of goods between member states is collected by the individual countries and reported to the EU Intrastat system. If the EU were to extend this data to all goods and services and merge it with the domestic VAT data, this could result in the first comprehensive multi-country firm-level supply network, representing nearly 20% of world GDP. In fact, the recent proposal "VAT in the digital age" (ViDA) by the European Commission could provide the legal framework and set a new international standard.

The EU could even further enlarge its map of supply dependencies to partially include non-member countries by merging its supply network data with customs data. Customs data features excellent coverage and detail on international firm-level trade and has already been linked with domestic VAT-based supply networks (*10*). The supply network of EU firms and their linkages to non-EU trading partners would yield a granular view into supply chain dependencies covering almost 40% of global trade.

A strategy to reconstruct the EU firm-level supply network could provide a blueprint for other economic and trade communities like Mercosur, the East African Union, or the Association of Southeast Asian Nations. Various Latin American, African, and Asian countries, including China and several states of India, already have the relevant data collection procedures in place. One of the major countries that does not collect VAT is the U.S. However, the U.S. is fairly well-covered in most commercial datasets and has extensive commodity flow surveys. The U.S. is already actively developing supply chain maps and considering the creation of a centralized supply chain data hub.

Once the supply networks for different economic blocs are mapped, the next step would be to connect these, again, via customs and trade data. To achieve a coherent database of international supply linkages, harmonized standards on data collection and formatting should be developed, e.g., by building on some of the best practices provided by individual countries. Just like what happened for the development of national accounts, such standardization efforts could be coordinated by international organizations like the UN, the International Monetary Fund (IMF), the World Bank, or the Organisation for Economic Co-operation and Development (OECD). A big advantage of internationally harmonized economic data, such as national accounts, is that they can be matched with other relevant datasets, enabling us to analyze the interaction of sectoral input-output dependencies with financial and environmental accounts or occupational statistics. A coordinated international effort could establish such a standard at the firm level.

However, since detailed supply chain data is highly sensitive, it will not be shared between regions with strained economic and political relationships in the foreseeable future. Thus, it is critical to develop suitable strategies for filling in the blind spots. Commercial datasets paired with network reconstruction methods can be used for reconstructing critical international supply chain dependencies. In recent years, much progress has been made in reconstructing missing information in complex networks, but these methods need to be carefully adapted to the specific context of supply connections.

Reconstruction methods for supply networks involve a mix of economic theory, careful accounting, combining public and commercial datasets, and state-of-the-art statistical methods. For example, we know a firm is more likely to buy from another firm if they are geographically close and have complementary product portfolios. If we observe enough interesting features of firms, we can train machine learning algorithms to recognize pairs of firms that are likely to trade. Other approaches have built on natural language processing, knowledge graphs, maximum entropy models, and leveraging telecommunication data (e.g., *3-11*, *12*, *13*). As in the past, these and new methods will evolve as more data becomes available.

A STRONG ALLIANCE

Drawing a trusted and comprehensive picture of international supply linkages that can effectively be used for policymaking requires the integration of multiple datasets, developing analytical tools, and establishing secure infrastructure for storing and processing sensitive information. Advancing this agenda requires a strong international alliance of various stakeholders, including national governments, statistical institutes, international organizations, central banks, the private sector, and the scientific community.

Supply chain data can be weaponized if it gets into the wrong hands (*14*), necessitating the highest standards for securely storing, sharing, and regulating the access and use of the data. Statistical offices have already developed protocols for providing controlled access while maintaining strict data security and privacy protections for national supply networks and other types of sensitive data, e.g., individual-level data in microdata centers. The European Health Data Space is another example where highly sensitive personal medical data will be collected and made available for research and health policy in a trustworthy manner. Similar approaches should be adopted for international supply chain data to ensure the highest data security and privacy standards while enabling meaningful use for research and policy.

As primary users of this data, national governments would strongly benefit and must play a key role in data collection and administration. The lack of granular economic microdata frequently hinders targeted policy approaches, resulting in ineffective redistribution of public funds. Transparent and safe access to detailed supply network data will enable governments to design evidence-based policies with greater precision. For example, such data could be used for identifying indirect exposures to natural disasters or upstream supply chain bottlenecks for critical goods, such as medical supplies. In this case, governments could take resilience-enhancing measures like incentivizing companies to diversify their suppliers.

Central banks could similarly benefit, as extensive firm-level supply network data will considerably improve our understanding of inflation dynamics and how the propagation of economic shocks affects financial stability (e.g., *15*). Central banks are already at the forefront of working with this data and making it accessible for scientific economic analysis. As data availability and methodological development progress, they could integrate this knowledge into monetary policy and financial market supervision.

Currently, there are various independent efforts at both the national and supranational levels to chart specific parts of the global supply network. Notable examples include the EU proposal for a Directive on corporate sustainability due diligence, the U.S. Supply Chain Disruptions Task Force, and the U.K. Department of International Trade supply chains resilience framework. However, without a concerted approach, these maps will remain fragmented and inadequate for addressing critical societal challenges. International organizations, including the IMF, the World Bank, and the OECD, have considerable expertise in harmonizing international datasets. Thus, they should play a key role in scaling up these efforts to the international level.

Building this alliance would result in a comprehensive map of international firm-level supply connections. This map could serve as a foundation for economic analyses and policies on both the national and international levels.

REFERENCES AND NOTES

1. A. Bacilieri, A. Borsos, P. Astudillo-Estevez, F. Lafond, Firm-level production networks: what do we (really) know?, Tech. rep., INET Oxford Working Paper No. 2023-08 (2023).
2. O. Celasun, N.-J. Hansen, A. Mineshima, M. Spector, J. Zhou, Supply bottlenecks: Where, why, how much, and what next?, Tech. rep., IMF Working Paper WP/22/31 (2022).
3. V. M. Carvalho, M. Nirei, Y. U. Saito, A. Tahbaz-Salehi, *The Quarterly Journal of Economics* **136**, 1255 (2020).
4. C. Diem, A. Borsos, T. Reisch, J. Kertész, S. Thurner, *Scientific reports* **12**, 1 (2022).
5. G. Culot, M. Podrecca, G. Nassimbeni, G. Orzes, M. Sartor, *Journal of Supply Chain Management* **59**, 1, (2023).
6. L. N. Ialongo *et al.*, *Scientific reports* **12**, 1 (2022).
7. T. Y. Choi, K. J. Dooley, M. Rungtusanatham, *Journal of Operations Management* **19**, 351 (2001).
8. E. Dhyne, G. Magerman, S. Rubínová, Tech. rep., NBB Working Paper 288 (2015).
9. D. Pomeranz, *American Economic Review* **105**, 2539 (2015).
10. E. Dhyne, A. K. Kikkawa, M. Mogstad, F. Tintelnot, *The Review of Economic Studies* **88**, 643 (2021).
11. A. Brintrup *et al.*, *Complexity* 9104387 (2018).
12. E. E. Kosasih, A. Brintrup, *International Journal of Production Research* **60**, 17 (2021).
13. L. Mungo, F. Lafond, P. Astudillo-Estévez, J. D. Farmer, *Journal of Economic Dynamics and Control* 104607 (2023).
14. H. Farrell, A. L. Newman, *Nature* **605**, 219 (2022).
15. C. Duprez, G. Magerman, Price updating in production networks, Tech. rep., NBB Working Paper 352 (2018).

ACKNOWLEDGMENTS

The authors thank L. Yang for the visualization, J. Yang for help with the data, and the participants at the *1st Interdisciplinary Workshop on Firm-Level Supply Networks in November 2022* at the Complexity Science Hub, Vienna, for stimulating discussions. A.P. received funding from the James S. McDonnell Foundation (2020-1640). C.D. received funding from the OeNB Anniversary Fund (P18696) and FFG (886360). F.L. received funding from Baillie Gifford and from the UKRI through ESRC grant PRINZ (ES/W010356/1). G.M. received funding from ARC grant (20204850) and FNRS (H.P046.20). V.M.C. received funding from ERC Consolidator Grant MICRO2MACRO (#GAP-101001221). J.D.F. received funding from Baillie Gifford. S.T. received funding from FFG (882184 and 886360).