Check for updates

The World Economy

WILEY

The World Economy

ORIGINAL ARTICLE OPEN ACCESS

Is the World Getting Smaller? Network-Driven Heterogeneity in Gravity-Model Elasticities

Zsolt T. Kosztyán[1,2,3] | Rita Mátrai[2] | Dénes Kiss[2]

[1]Faculty of Business and Economics, Business Research Institute for Data-Driven Growth and Economics, University of Pannonia, Veszprém, Hungary | [2]Department of Quantitative Methods, Faculty of Business and Economics, Institute of Management, University of Pannonia, Veszprém, Hungary | [3]Institute of Advanced Studies, Kőszeg (iASK), Kőszeg, Hungary

Correspondence: Zsolt T. Kosztyán (kosztyan.zsolt@gtk.uni-pannon.hu)

Received: 16 March 2026 | **Revised:** 17 July 2026 | **Accepted:** 13 August 2026

Keywords: extensive margin | fractal dimension | Granger causality | lacunarity | multilateral resistance | structural gravity | trade network topology

ABSTRACT

Gravity models constitute the workhorse of empirical international trade research, yet the temporal stability of their core parameters—distance and GDP elasticities—remains underexplored, particularly in relation to the evolving structure of trade networks. This study investigates the bidirectional Granger-causal relationships between gravity model coefficients and trade network topology using two complementary databases: the OECD Inter-Country Input–Output tables (1995–2020, 45 sectors) and the CEPII-BACI bilateral trade database (1995–2022, 97 product groups). We employ three methodologically distinct causality testing approaches—traditional panel Granger tests, Bayesian model comparison via Bayes factors, and Random Forest-based variable importance analysis—to establish robust directional relationships while accounting for potential nonlinearities. A key methodological innovation is the incorporation of fractal dimensions and local lacunarity distribution moments (mean, variance, skewness, kurtosis) as novel network indicators capturing multiscale heterogeneity, alongside conventional centrality and modularity measures. Using network-based dimensionality reduction (Generalized Network-based Dimensionality Analysis), we identify two to four distinct temporal clusters of sectors and products exhibiting heterogeneous gravity coefficient dynamics, with forecasts extended to 2030 using Bayesian ARIMA models. Our results reveal significant Granger-causal relationships between network structural properties and gravity coefficients, but the three methods diverge sharply in how much structure they detect: the Random Forest approach identifies three- to four-fold more significant sector-and product-level relationships than the two linear methods. At the aggregate level, centralization measures and heterogeneity indicators (lacunarity, fractal dimension) Granger-precede GDP elasticities broadly and consistently across sectors and products, whereas their relationship with distance elasticity is concentrated in specific sectors and product groups rather than operating uniformly across the aggregate network. Distance elasticity shows

Abbreviations: **AC**, authority centrality; **ALE**, average local efficiency; **ARIMA**, autoregressive integrated moving average; **Assort**, assortativity; **AVPL**, average path length; **BACI**, Base pour l'Analyse du Commerce International; **BC**, Betweenness Centrality; **BZ**, betweenness centralization; **CEPII**, Centre d'Etudes Prospectives et d'Informations Internationales; **CGE**, computable general equilibrium; **DCI**, in-degree centrality; **DCO**, out-degree centrality; **Dens**, density; **DZI**, in-degree centralization; **DZO**, out-degree centralization; **EC**, eigenvector centrality; **ECC**, eccentricity; **ERGM**, exponential random graph models; **ES**, Edge Similarity; **FDim**, fractal dimension; **GDP**, gross domestic product; **GLE**, global efficiency; **GNDA**, Generalized Network-based Dimensional Analysis; **GVC**, global value chain; **HC**, hub centrality; **HDLSS**, high dimensional low sample size; **HS**, harmonized system; **HZ**, harmonic centralization; **ICIO**, inter-country input–output table; **ICT**, infocommunication technology; **IP**, Intellectual Property; **ISIC**, International Standard Industrial Classification; **JZS**, Jeffreys–Zellner–Siow; **KC**, Katz Centrality; **KCr**, K-core; **LEV**, local efficiency value; **LL**, local lacunarity; **LV**, latent variable; **MCMC**, Markov Chain Monte Carlo; **ML**, machine learning; **MLE**, Maximum-Likehood Estimation; **MOD**, modularity value; **MOT**, Motif; **NS**, node similarity; **OECD**, Organization for Economic Co-operation and Development; **OLS**, ordinary least square; **PCA**, principal component analysis; **PD**, degree-degree pearson correlation; **PFA**, principal factor analysis; **PPML**, Pseudo Poisson Maximum Likelihood; **PRC**, PageRank centrality; **PRZ**, PageRank centralization; **PZ**, power centralization; **QAP**, quadratic assignment procedures; **RCC**, rich-club coefficient; **REC**, reciprocity; **RG**, research gap; **RQ**, research question; **RRes**, resilience (random attack); **SCI**, in-strength centrality; **SCO**, out-strength centrality; **SD**, degree-degree Spearman correlation; **SP**, shortest path similarity; **SRes**, resilience (systematic attack); **TERGM**, temporal exponential random graph models; **Trans**, transitivity; **WTN**, World Trade Network.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *The World Economy* published by John Wiley & Sons Ltd.

declining magnitude over the study period at the aggregate level, but substantial heterogeneity persists across product groups, with some clusters exhibiting stable or even increasing distance sensitivity. The cross-database triangulation confirms that these patterns are robust to alternative trade flow conceptualizations (gross vs. value-added trade). These findings contribute to international trade theory in three ways. First, they provide systematic evidence of temporal precedence—established via Granger-style tests, not structural causal identification—showing that the extensive margin of trade—captured by network topology—consistently leads the intensive margin reflected in gravity coefficients, consistent with heterogeneous firm models predicting that firms' entry/exit decisions reshape network structure before aggregate trade volumes adjust. Second, and most robustly, the documented sector-level heterogeneity offers a structural account of the 'distance puzzle': ICT and services sectors exhibit genuine distance compression through network decentralization, while heavy manufacturing maintains geographic constraints through persistent regional modularity. Third, the Granger-causal precedence of network fragility indicators over gravity coefficients establishes an empirical bridge between the production network literature on shock propagation and the gravity framework, with direct relevance for understanding supply chain resilience amid geopolitical fragmentation.
JEL Classification: F10, F14, C23, L14, D85

1 | Introduction

Gravity models serve as the fundamental framework for empirical international commerce, elucidating the relationship between bilateral trade and the economic size of countries (usually represented by gross domestic product (GDP)) as well as trade frictions, particularly geographic distance. In structural gravity, the income ('mass') elasticities and distance elasticity are not simply reduced-form coefficients; they represent fundamental characteristics and equilibrium entities associated with bilateral and multilateral trade costs, general equilibrium incidence, and market accessibility (Anderson and van Wincoop 2003; Fally 2015; Head and Mayer 2014). Nevertheless, the majority of practical research regards these coefficients as invariant within samples or across extensive periods, despite significant transformations in technology, policy, and global industrial organization since the mid-1990s.

A growing literature documents that gravity relationships are not invariant over time. The so-called distance puzzle—whether distance matters less in trade—has been debated for decades (Brun et al. 2005; Disdier and Head 2008; Berthelon and Freund 2008; Anderson and Yotov 2010; Novy 2013; Yotov 2012). Although Infocommunication Technology (ICT) diffusion and platforms can compress distance in some contexts (Lendle et al. 2016), compositional changes and quality differentiation can preserve or even amplify the role of distance in others (Berthelon and Freund 2008; Hummels and Klenow 2005; Baldwin and Harrigan 2011). More broadly, time variation in gravity coefficients can reflect evolving multilateral resistance, changing trade costs (tariff and non-tariff), supply-chain lengthening or reshoring, or shifts in the elasticity of substitution (Anderson and van Wincoop 2003; Head and Mayer 2014; Novy 2013; Eaton and Kortum 2002; Simonovska and Waugh 2014; Egger and Nigai 2015).

At the same time, world trade is a multilayer, evolving network of links among countries, sectors, and products (Fagiolo et al. 2009; Benedictis and Tajoli 2011; Barigozzi et al. 2010; Garlaschelli and Loffredo 2004; Cerina et al. 2015; Kivelä et al. 2014). Network structure—centrality, clustering, assortativity—both shapes and is shaped by trade frictions and market size (Chaney 2014; Opsahl et al. 2010). The fragmentation of production across borders global value chains (GVCs) further implies that sector- and product-specific gravity relationships may follow distinct temporal paths as value-added flows rewire the underlying network (Johnson and Noguera 2012; Koopman et al. 2014; Cerina et al. 2015; Antràs and Chor 2013). Ignoring such heterogeneity risks masking economically and statistically meaningful dynamics that matter for forecasting, counterfactual analysis, and policy design.

The benefits extend beyond fit. Theoretical and empirical research shows that economic networks condition resilience and shock propagation (Acemoglu et al. 2012; Carvalho 2014; di Giovanni et al. 2020; Boehm et al. 2019; Bonadio et al. 2021; Carvalho et al. 2021; Vespignani 2012). If gravity coefficients track the evolving strength of trade frictions and market-access elasticities, their temporal co-movement with network structure can reveal which layers of the trade network (countries, sectors, products) are fragile or robust to shocks, where bottlenecks arise, and how shocks propagate internationally.

The intersection of gravity models and network analysis in international trade research has garnered increasing attention over the past decade, yet remains incompletely explored. Although gravity models have long served as the workhorse for understanding bilateral trade flows based on economic mass and geographic distance (Anderson and van Wincoop 2003), network science has emerged as a complementary framework for capturing the structural properties and evolutionary dynamics of global trade systems (Sgrignoli et al. 2015). Lee et al. (2025) proposed a data-driven method to infer economic mass directly from trade flows, capturing intrinsic trade capacity as a node-level network attribute; Flori and Spelta (2025) introduced a community detection framework based on gravity models to analyse the mesoscale structure of the carbon trade network; Robinson et al. (2025) derived the gravity model from first principles using an electrical circuit analogy, naturally producing core-periphery structures observed in real trade networks. However, these studies predominantly focus on static or short-term analyses (He et al. 2024; Fang et al. 2023; Jiang et al. 2023), examining either the application of gravity models to explain network formation or using network indicators as covariates in gravity regressions, without systematically investigating the temporal precedence and Granger-causal relationships between network topology and gravity model parameters using multiple complementary methods.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Critical research gaps remain unaddressed in the literature. First, the mechanism of action—whether network structural changes drive shifts in gravity coefficients or vice versa—has not been subjected to rigorous causality analysis. Although several studies employ temporal exponential random graph models (TERGM) (Ding et al. 2025; Feng et al. 2024) or Quadratic Assignment Procedures (QAP) (Xing and Chen 2023; Dong et al. 2024) to identify determinants of network evolution, these methods primarily test the influence of exogenous factors on network formation rather than the dynamic interplay between endogenous network topology (e.g., fractal dimensions, lacunarity) and gravity model parameters. Second, existing network explorations are incomplete: most studies focus on conventional indicators such as degree centrality, betweenness, and modularity (Hao et al. 2020; García-Perez et al. 2016; Chen et al. 2023), while advanced topological measures like local lacunarity and network box-counting fractal dimensions—which capture multiscale heterogeneity and self-similarity—remain unexplored in the gravity model context. Third, research has been fragmented across specific sectors (e.g., virtual water trade (Xing and Chen 2023), waste trade (Poon et al. 2024), renewable energy (Feng et al. 2024)) or regions, lacking comprehensive studies that examine multiple trade databases (Organization for Economic Co-operation and Development (OECD) Inter-Country Input–Output table (ICIO) vs. Centre d'Etudes Prospectives et d'Informations Internationales (CEPII)-Base pour l'Analyse du Commerce International (BACI)) over extended periods (1995–2022) to assess the robustness and generalizability of network-gravity relationships. Our research addresses these Research Gaps (RGs) by

- **RG_1** conducting traditional, Bayesian, and Machine Learning (ML)-based panel Granger causality analyses to establish directional temporal relationships between network topology and gravity coefficients.
- **RG_2** incorporating fractal dimensions and local lacunarity distribution moments (mean, standard deviation, skewness, kurtosis) as novel network indicators that capture multiscale heterogeneity.
- **RG_3** performing comparative temporal analysis across two major trade databases (OECD ICIO and CEPII-BACI) spanning 1995–2022, enabling cross-validation at sectoral and product levels.
- **RG_4** investigating how changes in network structural properties—including modularity, centralization measures, and fractal characteristics—temporally precede or follow shifts in distance sensitivity and GDP elasticities.

A comprehensive study examining temporal relationships across different trade networks is essential for several reasons. First, different databases capture distinct dimensions of trade: OECD ICIO emphasizes value-added trade and sectoral linkages, while CEPII-BACI provides high-resolution product-level bilateral flows. Cross-database validation ensures that observed network-gravity relationships are not artefacts of data construction methodologies but reflect genuine economic phenomena. Second, temporal analysis reveals whether the relationships between network indicators and gravity coefficients are stable or exhibit regime shifts—critical knowledge for building reliable forecasting models. Third, identifying whether network structural properties can serve as leading indicators for changes in distance sensitivity would enable more proactive trade policy design.

We leverage two complementary sources—the OECD-ICIO framework and the CEPII-BACI product-level database—to estimate and analyse time-varying gravity coefficients from 1995 to 2022 at multiple levels of aggregation. This creates a unique opportunity to relate structural gravity dynamics to network evolution across sectors and products using consistent, high-quality data (Gaulier and Zignago 2010; Mayer and Zignago 2011; Timmer et al. 2015).

Based on the research gaps we formulated the following Research Question (RQ): How do the income and distance coefficients of structural gravity evolve over 1995–2022 at the network-wide, sectoral (OECD-ICIO), and product-group (CEPII-BACI) levels? Do changes in trade-network structure (centrality, modularity, fractal dimensions, lacunarity) Granger-cause shifts in these gravity coefficients, or vice versa? What patterns of heterogeneity emerge across sectors and products, and what do these dynamics imply for understanding trade network resilience?

The remainder of this paper is organized as follows. Section 3 reviews the relevant literature on structural gravity models, trade networks, time-varying coefficients, and causality methods, identifying the specific gaps our research addresses. Section 4 describes the two databases employed (OECD-ICIO and CEPII-BACI), the construction of multiplex trade networks, the network indicators calculated, and the three causality testing methodologies applied. Section 5 presents the empirical findings, including temporal patterns of gravity coefficients, network indicator evolution, correlation analyses, and causality test results across aggregate, sectoral, and product-group levels. Section 6 interprets these findings in light of existing theory, discusses the mechanisms connecting network structure to gravity parameters, and examines the heterogeneity across sectors and products. Section 7 summarizes the main contributions, discusses methodological innovations and policy implications, and outlines directions for future research. Technical details on gravity model estimation, network indicator calculations, forecasting methods, and causality testing procedures are provided in the Appendix A.

2 | Theoretical Framework

Our analysis is grounded in three interconnected theoretical strands that jointly motivate the hypothesized relationships between trade network topology and gravity model parameters. This framework connects the structural gravity tradition (Anderson and van Wincoop 2003; Head and Mayer 2014; Yotov 2024) with heterogeneous firm trade theory (Melitz and Redding 2015; Chaney 2014) and the production network literature (Acemoglu et al. 2012; Baqaee and Farhi 2019).

2.1 | Structural Gravity and the Economics of Distance Elasticity

The structural gravity model, derived from the seminal contributions of Anderson (1979), Anderson and van Wincoop (2003) and extended by Eaton and Kortum (2002), Helpman et al. (2008),

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and Melitz and Redding (2015), provides the micro-founded basis for our gravity specifications. In this framework, the distance elasticity γ_t is not merely a statistical coefficient but reflects the interaction of bilateral trade costs, the elasticity of substitution between varieties, and—in heterogeneous firm models—the shape of the productivity distribution (Carrère et al. 2020; Arkolakis et al. 2012).

Two properties of structural gravity are central to our analysis. First, Arkolakis et al. (2012) demonstrated that aggregate welfare gains from trade can be summarized by two sufficient statistics: the domestic expenditure share and the trade elasticity. As the distance elasticity γ_t captures a key component of the trade elasticity, temporal shifts in γ_t have direct welfare implications: a declining $|\gamma_t|$ implies increasing gains from trade, while a rising $|\gamma_t|$ signals growing trade frictions with potential welfare losses.

Second, the recent comprehensive surveys by Yotov (2024) and Larch et al. (2025) establish that structural gravity has evolved from an a-theoretical estimation equation to a full computable general equilibrium (CGE) framework. Our contribution leverages this evolution: by treating year-specific gravity estimates as realizations of a time-varying structural process, we investigate the economic forces—captured by network topology—that drive temporal shifts in these structural parameters.

2.2 | Network Topology as Extensive Margin Dynamics

Heterogeneous firm trade models predict that trade liberalization operates through two channels: the *intensive margin* (existing firms trade more) and the *extensive margin* (new firms enter export markets or existing firms exit) (Helpman et al. 2008; Melitz and Redding 2015). Chaney (2014) proved theoretically that firm-level heterogeneity in productivity and trade costs endogenously generates network structures with specific topological properties: fat-tailed degree distributions, high clustering, and core-periphery architecture.

We extend these insights by hypothesizing that *changes* in network topology—measured by centrality, modularity, fractal dimension, and lacunarity—primarily reflect extensive margin adjustments: firms entering new markets, forming new trade partnerships, or reorganizing supply chains. If these network-level changes are driven by the same trade cost fundamentals that determine gravity coefficients, but adjust at different speeds (because the extensive margin responds faster than the intensive margin), then network topology should exhibit temporal precedence over gravity parameters. This yields three testable hypotheses:

- **H_1 (Centralization → GDP elasticities):** Changes in network centralization and density Granger-cause shifts in GDP elasticities, because new trade links (extensive margin) restructure market access before aggregate flow volumes (intensive margin) fully adjust. Economically, this reflects the 'preferential attachment' mechanism: as hub economies attract new trade connections, their gravitational pull (GDP elasticity) strengthens with a lag of one year in traditional Granger tests and of four to five years once we require the effect to clear a Random-Forest shadow-feature significance threshold (Section Random Forest-Based Granger-Like Causality Analysis).
- **H_2 (Community structure → distance elasticity, heterogeneously across sectors):** Changes in network community structure Granger-cause shifts in distance elasticity, because regional trade bloc formation—captured by community structure (Caliendo and Parro 2015)—reduces within-bloc trade frictions with a lag as firms exploit new preferential access. At the aggregate network level this relationship is weak: of our 29 aggregate network indicators, only path-length (Average Path Length (AVPL)) and two centrality-adjacent measures Granger-precede the aggregate distance elasticity, and modularity itself does not. We therefore test H_2 primarily at the sectoral and product-group level (Figures 9 and 10), where the 'distance puzzle' (Disdier and Head 2008) should be attributable to network reorganization in some, but not all, sectors.
- **H_3 (Network heterogeneity → GDP elasticities, and more weakly to distance elasticity):** Multiscale heterogeneity measures (lacunarity, fractal dimension) Granger-cause changes in gravity coefficients, because structural irregularities in trade partner distributions reflect vulnerabilities that propagate through the input–output cascade mechanism (Acemoglu et al. 2012; Baqaee and Farhi 2019). At the aggregate level this holds clearly for GDP elasticities—higher lacunarity and lower fractal dimension precede subsequent changes in both origin and destination GDP elasticity—but not for the aggregate distance elasticity, where none of the four lacunarity moments or fractal dimension pass either the traditional or the Random-Forest shadow-feature test; as with H_2, any distance-elasticity effect of network heterogeneity appears to be concentrated in specific sectors and products rather than operating uniformly across the aggregate network (Section 6.5.3).

2.2.1 | Economic Interpretation of Fractal Dimension and Lacunarity

These two measures are not purely descriptive topological statistics; each maps onto a specific economic mechanism already implicit in the heterogeneous-firm and production-network literatures we build on. Following the network box-counting approach of Song et al. (2005), the fractal dimension of a trade network summarizes how self-similarly trade linkages are distributed across spatial/relational scales. A lower fractal dimension corresponds to a more hub-and-spoke, hierarchical topology in which a small number of intermediating hub economies account for a disproportionate share of connections at every scale, consistent with Chaney (2014)'s prediction that heterogeneous-firm entry generates core-periphery structure; a higher fractal dimension indicates a more space-filling, decentralized pattern of direct bilateral linkages, consistent with a lower effective cost of establishing new bilateral trade relationships (the extensive margin). Lacunarity, in turn, measures the 'gappiness'—the heterogeneity in the size and spacing of gaps between a country's trade partners—of the network at a given scale (as reflected in our own node-level definition, $\mathrm{LL}_i = \langle r_i^2 \rangle / \langle r_i \rangle^2$, Section Calculated Network Properties); economically, high lacunarity indicates

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

that a country's trading partners are unevenly distributed in 'network space', concentrated in intensively-connected clusters separated by large gaps to other potential partners. This is precisely the kind of structural irregularity that Baqaee and Farhi (2019) show can amplify the aggregate effect of local shocks nonlinearly, because a shock hitting a densely-connected cluster propagates disproportionately within it before reaching sparsely-connected regions. Rising lacunarity is therefore a leading indicator of increased effective distance sensitivity not because geographic distance itself has changed, but because uneven partner spacing raises the effective (network) distance a shock—or a new trade relationship—must travel, a mechanism related to, but distinct from, the geographic distance term already present in the gravity equation. We treat fractal dimension and lacunarity as summarizing multiscale structural heterogeneity that is complementary to, rather than a restatement of, conventional centrality and modularity measures, consistent with the only moderate correlations we document between FDim/LL and degree/eigenvector centrality. In line with H_3, the same logic extends this mechanism to the GDP elasticities that H_3 also predicts: a more space-filling (higher fractal dimension) or more evenly-spaced (lower lacunarity) topology reflects the same extensive-margin process of new bilateral-relationship formation already invoked for centralization under H_1, so fractal dimension and lacunarity should Granger-precede changes in both origin- and destination-GDP elasticities through this market-access channel, not only the distance elasticity discussed above. Unlike the distance channel, this market-access channel does not imply an unambiguous sign ex ante, so we treat the direction of this relationship for the GDP elasticities as an open empirical question, addressed directly by the GDP elasticity causality results in Section 5 and revisited in Section 6.

2.3 | Production Networks, Shock Propagation, and the Gravity–Resilience Nexus

The production network literature establishes that microeconomic shocks propagate through input–output linkages and can generate aggregate fluctuations (Acemoglu et al. 2012; Baqaee and Farhi 2019; Carvalho 2014). Baqaee and Farhi (2019) demonstrate that this propagation is inherently *nonlinear*: in networks with heterogeneous firms and complementary inputs, small shocks can have disproportionately large aggregate effects. This nonlinearity provides a theoretical basis for our use of Random Forest-based causality testing alongside linear methods.

Direct empirical evidence comes from Boehm et al. (2019) (natural disaster propagation) and Bonadio et al. (2021) (COVID-19 supply chain transmission). These studies document that network structure—particularly the concentration of trade through hub nodes—determines the speed and magnitude of shock propagation.

Our framework connects this literature to gravity models by proposing that network resilience indicators—particularly the divergence between random (RRes) and systematic (SRes) attack resilience—capture the structural vulnerability of trade networks to disruptions. When SRes deteriorates relative to RRes, the network becomes increasingly dependent on a few critical hub nodes. Such concentration should, with a lag, affect the effective trade cost structure captured by gravity coefficients because disruptions to hub countries propagate disproportionately through the network (Carvalho et al. 2021). This provides a theoretical mechanism for linking network resilience to gravity coefficient dynamics and connects our analysis to the growing literature on geopolitical fragmentation and supply chain reorganization (Alfaro and Chor 2023; Freund et al. 2024).

3 | Background

3.1 | Structural Gravity Models and Temporal Variation

The modern structural gravity model provides a theoretical foundation for gravity coefficients and multilateral resistance terms (Anderson 1979, 2011; Anderson and van Wincoop 2003; Head and Mayer 2014; Fally 2015). Advances in estimation address heteroskedasticity and zeros (Pseudo Poisson Maximum Likelihood (PPML)) and advocate high-dimensional fixed effects to capture multilateral resistance and time-varying shocks (Santos Silva and Tenreyro 2006; Helpman et al. 2008; Silva and Tenreyro 2011; Head and Mayer 2014; Fally 2015). Against this background, the temporal stability of gravity coefficients has been scrutinized via the distance puzzle and related decompositions (Brun et al. 2005; Disdier and Head 2008; Berthelon and Freund 2008; Anderson and Yotov 2010; Yotov 2012; Novy 2013), with mixed evidence: ICT-enabled platforms can attenuate distance (Lendle et al. 2016), whereas composition and quality can sustain or magnify it (Berthelon and Freund 2008; Hummels and Klenow 2005; Baldwin and Harrigan 2011). Micro-founded models of trade costs and elasticities provide additional interpretations for evolving coefficients (Eaton and Kortum 2002; Simonovska and Waugh 2014; Egger and Nigai 2015). Methodological guidance emphasizes correct specification in panels and the treatment of bilateral dependence (Egger and Pfaffermayr 2003; Baldwin and Taglioni 2006; Head and Mayer 2014; Pesaran 2006; Chudik and Pesaran 2015).

Recent surveys by Yotov (2024) and Larch et al. (2025) have comprehensively documented the evolution of structural gravity from an a-theoretical estimating equation to a full general equilibrium framework. Yotov (2024) emphasizes that proper treatment of multilateral resistance terms, zero trade flows, and country-pair heterogeneity is essential for obtaining unbiased gravity estimates. Larch et al. (2025) provide specific practical recommendations, including the use of PPML estimation (Santos Silva and Tenreyro 2006) and country-pair fixed effects. Our cross-sectional approach, while departing from the panel fixed-effects standard, is shared with the distance puzzle literature (Disdier and Head 2008; Yotov 2012) and is necessary for our research design, which requires year-specific coefficient estimates. We address potential biases through PPML robustness checks (Section Robustness Check) and cross-database validation.

Antràs (2025) provides a comprehensive overview of the frontier in international trade theory, highlighting the importance of heterogeneous firms, global sourcing, and network effects for understanding modern trade patterns. Our analysis directly addresses his call for empirical work connecting network-level phenomena to structural trade parameters.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3.2 | Identification Challenges in Structural Gravity

A critical strand of the gravity literature has focused on identification and estimation challenges that bear directly on the interpretation of time-varying coefficients. We discuss four issues particularly relevant for our analysis, following the comprehensive guides by Head and Mayer (2014), Yotov (2024), and Larch et al. (2025).

3.2.1 | First, Multilateral Resistance

Anderson and van Wincoop (2003) demonstrated that bilateral trade depends not only on bilateral trade costs but also on trade costs with all other partners—the multilateral resistance terms (MRTs). Failing to account for MRTs produces biased and inconsistent estimates of distance and GDP elasticities (Head and Mayer 2014; Yotov 2024). The standard remedy is exporter- and importer-specific fixed effects, but, as explained in Section 4.1.3, these cannot be added to our year-by-year cross-sectional specification (1) without absorbing $\ln(GDP_{i,t})$ or $\ln(GDP_{j,t})$ themselves: within a single year, each exporter (importer) has exactly one GDP value, so the fixed effect and the corresponding GDP elasticity would be perfectly collinear. We instead include nine bilateral controls from the CEPII Gravity database (Conte et al. 2022) (Section 4.1.3), which absorb time-invariant and slow-moving *bilateral* heterogeneity but cannot substitute for genuinely *multilateral* (country-specific, all-partners) resistance terms.

Unlike a specification with exporter-/importer-year fixed effects, our design therefore does not remove within-year multilateral resistance from the dependent variable *by construction*. Our distance and GDP elasticities remain potentially exposed to time-varying MRTs that are correlated with network topology: any Granger-causal link we detect from network topology to the residual gravity coefficients would be *consistent with* network structure proxying for unmodeled MRT dynamics rather than exerting an independent effect, and we cannot rule this out with the current design. We flag this explicitly here and in Section 8, and, for this reason, we describe our findings throughout using temporal-precedence rather than causal language (see our response to the reviewer's first point), and return to this channel in the Discussion (Section 6.6).

3.2.2 | Second, Endogeneity of Trade Costs

Trade policy variables (tariffs, FTAs) are potentially endogenous: countries self-select into agreements based on anticipated trade gains (Baier and Bergstrand 2007; Larch et al. 2025). Equation (1) includes a joint FTA/WTO-membership indicator among its nine CEPII bilateral controls (Conte et al. 2022), together with contiguity, common official and ethnic language, colonial history, common legal origin, and EU membership of either partner (Section 4.1.3); several of the omitted-bilateral-factor concerns raised in the distance puzzle literature (Disdier and Head 2008; Brun et al. 2005; Yotov 2012) are therefore addressed directly rather than left in the error term. This does not resolve the deeper endogeneity concern that FTA participation is itself a policy choice correlated with unobserved bilateral trade potential, nor does it capture the country-specific policy environments that exporter-/importer-year fixed effects would otherwise absorb (see *First*, above); our distance elasticity should accordingly be interpreted as a partial, not fully causal, effect of geographic distance net of the observed bilateral controls.

3.2.3 | Third, Zero Trade Flows

The log-linear gravity model drops zero bilateral flows, potentially introducing selection bias (Santos Silva and Tenreyro 2006; Helpman et al. 2008). We address this directly: throughout the paper, Equation (1) is estimated by PPML on trade in levels rather than by Ordinary Least Square (OLS) on log-transformed trade, which lets zero-flow observations enter the estimation naturally and is robust to heteroskedasticity (Santos Silva and Tenreyro 2006) (Section A.6). Section Comparison of OLS and PPML Estimates for the Bilateral Control Variables reports the corresponding OLS estimates for comparison and confirms that our PPML-based coefficient trajectories are not an artefact of the estimator choice.

3.2.4 | Fourth, the Role of Country-Pair Fixed Effects

The state-of-the-art in gravity estimation employs country-pair fixed effects in panel settings to absorb all time-invariant bilateral factors, including distance (Larch et al. 2025; Yotov 2024). Our research design deliberately departs from this approach: we *require* cross-sectional distance estimates for each year to construct the time series that are the object of our Granger causality analysis. This design choice is shared with the distance puzzle literature (Disdier and Head 2008; Brun et al. 2005; Yotov 2012), which similarly needs year-specific distance coefficients. Yotov (2012) demonstrated that the distance puzzle is partly an artefact of omitting domestic trade flows, while our focus is on the *temporal co-movement* between distance coefficients and network topology rather than the absolute level of distance elasticity. We mitigate omitted variable concerns through comprehensive exporter and importer fixed effects in each cross-section and cross-database validation (OECD-ICIO vs. BACI-CEPII).

Because distance itself is time-invariant for a given country pair, this absorption is not merely a stylistic choice but a mechanical identification constraint: a specification with country-pair fixed effects cannot simultaneously deliver an estimate of the distance elasticity, since $\ln Dist_{ij}$ would be perfectly collinear with the pair fixed effect and would be dropped from the regression. Our research question requires precisely the opposite—a *separate*, identified distance elasticity for every year—which is why, like the rest of the year-specific distance-puzzle literature (Disdier and Head 2008; Brun et al. 2005; Yotov 2012), we use exporter- and importer-year fixed effects rather than country-pair fixed effects.

3.3 | Trade Networks and Multilayer Structure

The World Trade Network (WTN) exhibits rich, evolving topology—heavy-tailed degree and weight distributions, clustering, and core-periphery features (Fagiolo et al. 2009; Garlaschelli

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and Loffredo 2004; Serrano and Boguñá 2003). Recent work has begun to explicitly link gravity model predictions with network topology, demonstrating that network structure both reflects and shapes bilateral trade costs (Hood and Gassner 2017). Network approaches highlight heterogeneity across layers (countries, sectors, products) and the role of centrality for aggregate outcomes (Benedictis and Tajoli 2011; Barigozzi et al. 2010; Cerina et al. 2015; Kivelä et al. 2014; Opsahl et al. 2010). Gravity-consistent mechanisms can generate network structure endogenously (e.g., firm-to-firm matching and intermediary roles) (Chaney 2014). The expansion of GVCs underscores the need to disaggregate by sector/product and consider value-added flows when connecting gravity to network metrics (Johnson and Noguera 2012; Koopman et al. 2014; Antràs and Chor 2013; Cerina et al. 2015). Product-space and related perspectives further link specialization to network proximity (Hidalgo and Hausmann 2009).

3.4 | Heterogeneous Firms and Endogenous Network Formation

A central theoretical insight motivating our analysis comes from the heterogeneous firm trade literature. Melitz and Redding (2015) demonstrate that trade liberalization generates welfare gains through a selection channel: the most productive firms enter export markets while the least productive exit. Arkolakis et al. (2012) show that, under certain conditions, aggregate welfare gains can be summarized by two sufficient statistics: the domestic expenditure share and the trade elasticity—precisely the objects our gravity model captures.

Critically, Chaney (2014) proved that heterogeneous firm models endogenously generate trade network structures with specific topological properties. The degree distribution, clustering coefficient, and network density are all determined by the underlying productivity distribution and bilateral trade costs. This implies a tight theoretical connection between gravity model parameters (reflecting trade costs and elasticities) and network topology (reflecting firm entry/exit patterns).

Bernard and Moxnes (2018) survey the growing literature on networks and trade, emphasizing that firm-level network formation is driven by search frictions, relationship-specific investments, and information costs distinct from iceberg trade costs. This suggests that network topology captures trade friction information not fully reflected in distance or GDP alone, providing an economic rationale for why network indicators add predictive content for gravity parameters.

3.5 | Time-Varying Gravity Coefficients

Several strands explicitly call for or document temporal variation in gravity parameters, particularly the distance elasticity and the incidence of geography (Brun et al. 2005; Disdier and Head 2008; Berthelon and Freund 2008; Anderson and Yotov 2010; Yotov 2012; Novy 2013; Head and Mayer 2014). These contributions argue that ignoring time variation can confound policy evaluation and mismeasure trade costs, especially when multilateral resistance and compositional effects shift over time. At finer granularity, sector/product heterogeneity motivates estimating and comparing time paths across layers (Baldwin and Harrigan 2011; Hummels and Klenow 2005; Egger and Nigai 2015).

Linking the time paths of gravity coefficients to evolving network structure can identify whether changes in trade frictions or market size elasticities are associated with re-wiring of trade links and re-weighting across sectors/products. This is critical for understanding resilience and international shock transmission—how disruptions in one sector/country propagate through trade linkages (Acemoglu et al. 2012; Carvalho 2014; di Giovanni et al. 2020; Boehm et al. 2019; Bonadio et al. 2021; Carvalho et al. 2021; Vespignani 2012). Despite vast gravity and network literatures, direct causal mapping between time-varying gravity coefficients and structural network metrics remains scarce, especially at product and sector levels with long panels. Our study addresses this gap using consistent estimation on OECD-ICIO sectors and CEPII-BACI products (Gaulier and Zignago 2010; Mayer and Zignago 2011; Timmer et al. 2015).

3.6 | Causality Methods for Short Panels

Many sector/product panels feature short time dimensions, necessitating appropriate econometric methods. Classical Granger causality (Granger 1969) provides the baseline framework; panel extensions (Dumitrescu and Hurlin 2012) handle cross-sectional units with fixed effects. Bayesian model comparison via Bayes factors (Kass and Raftery 1995) offers an alternative that naturally incorporates parameter uncertainty and avoids multiple testing issues. Complementing linear methods with ML-based approaches such as Random Forests (Breiman 2001) can capture nonlinear relationships that traditional tests may miss. Because time-series dependence invalidates naive cross-validation, blocked or rolling procedures are needed for predictive evaluation (Bergmeir et al. 2018).

3.7 | Network Resilience and Shock Propagation

Production and trade networks amplify or dampen shocks depending on topology and link weights (Acemoglu et al. 2012; Carvalho 2014; Vespignani 2012). Firm- and sector-level evidence shows how input linkages transmit shocks within and across borders (Boehm et al. 2019; di Giovanni et al. 2020; Carvalho et al. 2021; Bonadio et al. 2021). By relating time variation in gravity coefficients to changes in network metrics, one can identify when higher distance sensitivity or lower market-size elasticity coincides with more brittle structures (e.g., hub-dependence, low redundancy), informing policy about diversification, regionalization, or supply-chain risk. To our knowledge, few studies jointly model the dynamics of structural gravity coefficients and the co-evolution of sector/product-level network structure; doing so can bridge the structural gravity and production-network literatures.

The empirical literature on the intersection of gravity models and network analysis in trade has expanded considerably since the mid-2010s, driven by advances in computational methods

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and data availability. Early contributions focused primarily on applying gravity models to predict trade flows, treating network structure as an exogenous outcome rather than an interactive system component. More recent work has begun to recognize the bidirectional relationship between network topology and trade determinants, employing sophisticated statistical frameworks such as exponential random graph models (ERGM), Temporal ERGM (TERGM), and QAP to disentangle the complex interplay between structural and behavioural factors. However, as detailed in Table 1, the existing body of research exhibits important limitations: most studies are sector-specific or geographically constrained, few employ explicit causality testing beyond correlation analysis, and almost none investigate advanced topological indicators such as fractal dimensions or lacunarity.The synthesis of this literature reveals several key patterns. Most studies focus on partial trade networks examining specific sectors or products, with few analysing complete global networks over extended periods. While the field has evolved toward dynamic analysis using ERGM/TERGM frameworks and QAP models, true causal inference techniques—such as Granger causality tests—are conspicuously absent. The literature exhibits substantial overlap in basic indicators (degree centrality, betweenness, modularity) but lacks exploration of advanced topological measures like fractal dimensions or lacunarity, despite theoretical arguments that multiscale structural properties may fundamentally relate to gravity model parameters.

In comparison, our research advances the field in four critical dimensions. *First*, we conduct genuine temporal causality analysis using traditional, Bayesian, and ML-based Granger causality tests and related methods to establish whether network indicators lead, lag, or co-evolve with gravity coefficients (see C_1). *Second*, we introduce novel network indicators—fractal dimensions and local lacunarity distributions—that capture scale-free properties and spatial heterogeneity more comprehensively than conventional metrics (see C_2). *Third*, we perform parallel analyses on two major trade databases (OECD ICIO and BACI-CEPII) spanning 1995–2022, enabling robust cross-validation and comparison of value-added versus product-level trade patterns (see C_1). *Fourth*, we examine complete trade networks rather than sector-specific slices, providing a system-level perspective on how macrostructural changes relate to fundamental trade determinants (see C_3). These innovations position our study as the first comprehensive, methodologically rigorous investigation of bidirectional temporal relationships between network topology and gravity model parameters, addressing the critical gaps identified in the existing literature.

- **C_1** (Theoretical): We provide systematic evidence of temporal precedence—not structural causal identification—linking the *extensive margin* of trade (network topology formation) to the *intensive margin* (gravity coefficient magnitudes), bridging two previously disconnected literatures: the structural gravity framework (Anderson and van Wincoop 2003; Head and Mayer 2014; Yotov 2024) and the production network approach (Acemoglu et al. 2012; Chaney 2014; Bernard and Moxnes 2018). Our finding that network topology changes Granger-precede gravity shifts—but rarely vice versa—is consistent with heterogeneous firm models (Helpman et al. 2008; Melitz and Redding 2015) predicting that firm-level entry/exit decisions (reflected in topology) temporally precede aggregate trade pattern adjustments (reflected in gravity): as discussed in Section 8, common shocks to both series cannot be ruled out, so we treat this as suggestive rather than definitive support for the mechanism.
- **C_2** (Empirical): We document that the 'distance puzzle' (Disdier and Head 2008) is not a unitary phenomenon but a composite of heterogeneous, sector-specific dynamics: services and ICT sectors exhibit declining distance sensitivity consistent with digital trade theories (Lendle et al. 2016; Antràs 2025), while heavy manufacturing maintains stable geographic constraints aligned with transport cost models (Hummels and Klenow 2005). This heterogeneity has first-order implications for understanding global value chain resilience (Bonadio et al. 2021; Carvalho et al. 2021) and the differential impact of geopolitical fragmentation across sectors (Alfaro and Chor 2023; Freund et al. 2024).
- **C_3** (Methodological–Policy): We introduce a triangulated causality framework combining traditional, Bayesian, and machine learning approaches with novel network indicators (fractal dimensions, lacunarity moments) to establish that network topology serves as a 1–4 year leading indicator for gravity coefficient shifts. This provides actionable early warning capacity for policymakers monitoring trade friction evolution and supply chain vulnerability.

These contributions directly address the identified research gaps: C_1 responds to RG_3 and RG_4 by providing comprehensive temporal analysis across databases; C_2 addresses RG_1 and RG_2 through rigorous causality testing with novel indicators; and Network Resilience and Shock Propagation tackles RG_4 by documenting systematic heterogeneity patterns.

4 | Material and Methods

Figure 1 shows the flowchart of the research. The main flowchart outlines the various stages of the study.

4.1 | Databases Used and Developed

This section describes the two primary data sources used in our analysis and the construction of edge lists for multiplex network analysis. This corresponds to **Step 1 (Data Collection and Preparation)** and **Step 2 (Multiplex Network Construction)** of our research framework.

4.1.1 | OECD Inter-Country Input–Output Database (OECD-ICIO)

The OECD Inter-Country Input–Output (ICIO) database provides a comprehensive framework for analysing global value chains and inter-sectoral linkages across countries. From this database, we exclusively utilize the **transaction matrix Z**, which captures intermediate consumption flows between country-sector pairs.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | Literature review: Studies examining relationships between gravity models and network indicators.

Authors	Database	Period	Network indicators	Scope	Static/ dynamic	Causality?	Key findings
Lee et al. (2025)	Trade flows (unspecified)	N/S	Node mass, intrinsic trade capacity	Complete	N/S	Yes	Developed a self-consistent method to infer economic mass directly from observed trade flows, eliminating reliance on GDP proxies. The approach successfully distinguished each country's intrinsic trade capacity from external factors, improving prediction accuracy
Flori and Spelta (2025)	Carbon trade	Recent phases	Mesoscale, community structure	Partial (carbon)	N/S	Yes	Applied a gravity model combined with optimal transport theory to detect communities in carbon allowance trading. Found that trades cluster within countries and sectors, with community patterns aligning with cultural dimensions rather than pure economic efficiency
Robinson et al. (2025)	Synthetic + real trade	N/S	Core-periphery, network-flow relations, friction	Complete	Dynamic	Yes	Derived the gravity model from first principles using an electrical circuit analogy where price differences drive flows against resistance (friction). The model naturally generates core-periphery structures and implies that trade flows self-organize to minimize total system friction
Xing and Chen (2023)	Multi-regional I-O	2000–2015	Path dependence, reciprocity, transmissive links	Partial (virtual water, 62 countries)	Dynamic	Yes	Endogenous network structures (reciprocity, path dependence, transmissive links) strongly explain virtual water trade network evolution. Traditional gravity variables (distance, GDP) showed partial support, but national water scarcity levels had no significant impact
Poon et al. (2024)	Waste trade	N/S	Multilayer, interlayer connections, hubs/havens	Partial (industrial/ textile waste)	N/S	Yes	Integrated multilayer network analysis with gravity models to test the Pollution Haven Hypothesis. Stringent environmental policies reduce both exports and imports. PHH was not supported for slag/dross but evident in textile waste trade between connected hubs and havens
Jiang et al. (2023)	Env. services	2001–2019	Core-edge, accessibility, convenience	Partial (env. services)	Dynamic	Yes (QAP)	Identified a clear core-edge structure with Belgium, Italy, and Netherlands at the center, and Greece emerging as a key bridge country. Climate change, geographic distance, and population size significantly affect trade, while economic distance and green technology distance show no significant effects

(Continues)

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | (Continued)

Authors	Database	Period	Network indicators	Scope	Static/ dynamic	Causality?	Key findings
He et al. (2024)	Environemental Goods	2002–2019	Core-periphery, density, product complexity	Partial (env. goods)	Dynamic	Yes	Demonstrated that technology complementarity promotes environmental goods trade more than technology competition, especially for high-complexity products. Trade patterns exhibit geographical regionalization, with complementarity driving North–North trade and competition promoting South–North exports of low-complexity products
Fang et al. (2023)	UN Comtrade	N/S	Small-world, scale-free, core-periphery, centrality	Partial (ginseng)	Dynamic	Yes (QAP)	Ginseng trade networks exhibit small-world and scale-free properties with core countries (China, South Korea, Germany, US) controlling trade. Geographic distance and GDP per capita negatively affect trade, while land adjacency, technological similarity, population size, and institutional similarity positively influence ginseng flows
Wang and Cui (2025)	Intellectual Property (IP) trade	N/S	Trade relationship formation	Partial (IP)	Dynamic	Yes	Smaller institutional distances between countries significantly increase the probability of intellectual property trade relationships forming. This effect persists after controlling for traditional gravity variables like geographic adjacency, common language, and colonial history
Feng et al. (2024)	Renewable energy (bilateral)	2000–2018	Small-world, reciprocity, assortativity, community	Partial (renewable)	Dynamic	Yes	Renewable energy trade networks show small-world properties and reciprocity patterns, with strong community structure. Endogenous network mechanisms (reciprocity, structural embeddedness) combine with gravity factors (continent, WTO, borders, language, religion, trade agreements) to determine trade patterns
García-Perez et al. (2016)	World Trade Atlas	1870–2013	Trade distances, hierarchization, hyperbolic architecture	Complete	Dynamic	Yes	Trade distances derived from gravity models successfully predict trade channel existence over 143 years. The distance coefficient grows over time and corresponds to the fractal dimension of the trade network, resolving the 'globalization puzzle'—networks become more hierarchical as distance matters more

(Continues)

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | (Continued)

Authors	Database	Period	Network indicators	Scope	Static/ dynamic	Causality?	Key findings
Hao et al. (2020)	Dutch broiler supply chain	N/S	Degree, betweenness, density, critical nodes	Partial (supply chain)	N/S	Yes	Used gravity models to estimate missing trade links in the Dutch broiler meat supply chain, then computed network centrality measures. Identified meat processors with their own slaughtering facilities as the most critical nodes for potential contamination spread
Kosowska-Stamirowska et al. (2016)	Maritime fleet movements	1977–2008	Common neighbours, system evolution, link forecasting	Partial (maritime)	Dynamic	Yes	Network effects, particularly the number of common neighbours between ports, govern maritime trade evolution in ways that improve upon standard gravity model assumptions. This simple network rule remains consistent over three decades and helps forecast new shipping line openings
Fagiolo et al. (2009)	Migration + trade	1960–2000	Topological structure, network centrality, correlations	Partial (migration/ trade)	Dynamic	Yes	A parsimonious gravity model successfully explains most of the international migration network topology. Countries more central in the migration network trade more with each other, with this correlation largely explained by economic size and geographic distance
Chung et al. (2020)	Meat trade (134 countries)	1995–2015	Structural changes, consolidated networks	Partial (meat)	Dynamic	Yes	Combined network modelling with cluster analysis to track structural changes in global meat trade over two decades. In consolidated networks, the presence of trade agreements and shorter geographic distances are strongly associated with increased bilateral meat trade volumes
Dong et al. (2024)	282 Chinese cities	2006–2019	Core-edge, community, microstructure, centrality	Partial (logistics carbon, China)	Dynamic	Yes (ERGM + QAP)	Applied modified gravity models and ERGM to analyse spatial correlation patterns in logistics carbon emission efficiency across Chinese cities. Financial investment, government concern, international openness, population density, and innovation ability all facilitate spatial correlation formation in the network
Chen et al. (2023)	Photovoltaic trade (bilateral)	2000–2019	Small-world, disassortativity, reciprocity, transitivity, stability	Partial (photovoltaic)	Dynamic	Yes (ERGM)	Photovoltaic trade networks exhibit small-world, disassortative, and low reciprocity characteristics. Network formation is driven by transitivity, reciprocity, and stability, combined with gravity factors including WTO membership, continental proximity, asymmetrical development levels, and regional trade agreements

(Continues)

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | (Continued)

Authors	Database	Period	Network indicators	Scope	Static/dynamic	Causality?	Key findings
Dueñas and Mandel (2023)	Music videos (GDELT implied)	N/S	Topology, modularity, asymmetric links, diffusion channels	Partial (cultural diffusion)	N/S	Yes	Applied gravity models to identify factors shaping global music diffusion networks. Found that cultural, geographic, and historical factors are primary drivers of network formation, while economic factors play a surprisingly minor role in cultural trade patterns
Ryczkowski et al. (2017)	Air transport	N/S	Hidden subflows, distance coefficients	Partial (air transport)	N/S	Yes	Developed a model for transfer flights that exploits discrepancies in standard gravity models to discover hidden subflows in the global passenger air-transport network. Successfully retrieved the gravity distance coefficient and suggested its correspondence to network fractal dimension

The transaction matrix **Z** is organized as a block matrix where each element z_{ij}^{rs} represents the value of intermediate goods and services flowing from sector *r* in country *i* to sector *s* in country *j*. The database covers: (1) *Time period:* 1995–2020 (annual data); (2) *Countries:* 76 economies (OECD members and major non-OECD economies); (3) *Sectors:* 45 industries based on International Standard Industrial Classification (ISIC) Rev. 4 classification.

From the transaction matrix **Z**, we construct a directed, weighted multiplex network where: (1) *nodes* represent countries; (2) *layers* represent economic sectors; (3) *edges* capture trade flows between countries, differentiated by *origin* and *destination* sectors; (4) *inter-layer connections* are explicitly modelled, representing cross-sectoral value chain linkages.

The edge list structure is defined in Table 2.

As quantities are not always interpretable, we worked with volume in all cases. The list of sectors is detailed in Table A5.

The following example presents an edge list format:

2019 | DEU | C29 | USA | C30 | 4523.7 | –

In this example, Germany's motor vehicle sector (C29) exports $4523.7 million worth of intermediate goods to the US transport equipment sector (C30) in 2019.

The presence of different origin and destination layers ($\texttt{Layer}_i \neq \texttt{Layer}_j$) explicitly captures inter-layer connections, which is a distinguishing feature of the OECD-ICIO multiplex network.

4.1.2 | BACI-CEPII International Trade Database

The BACI database, maintained by CEPII, provides harmonized bilateral trade flows at the product level using the Harmonized System (HS) classification. The list of product groups is listed in Table A6.

The BACI database contains: (1) *time period:* 1995–2022 (annual data); (2) *countries:* Over 200 reporting economies. (3) *products:* 6-digit HS codes (approximately 5000 product categories); (4) *volume:* trade values (thousands USD) and (5) *quantities:* in metric tons.

We construct three distinct multiplex networks based on different product aggregation levels:

1. **Product Group Level (HS-2):** First two characters of HS code (approximately 97 product groups).
2. **Product Subgroup Level (HS-4):** First four characters of HS code (approximately 1200 subgroups).
3. **Product Level (HS-6):** Full six-digit HS code (approximately 5000 products).

For the BACI-CEPII networks: (1) *nodes* represent countries; (2) *layers* represent products (at the chosen aggregation level); (3) *edges* capture bilateral trade flows.

Inter-layer connections are not defined, as products are treated as independent layers without explicit cross-product linkages.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

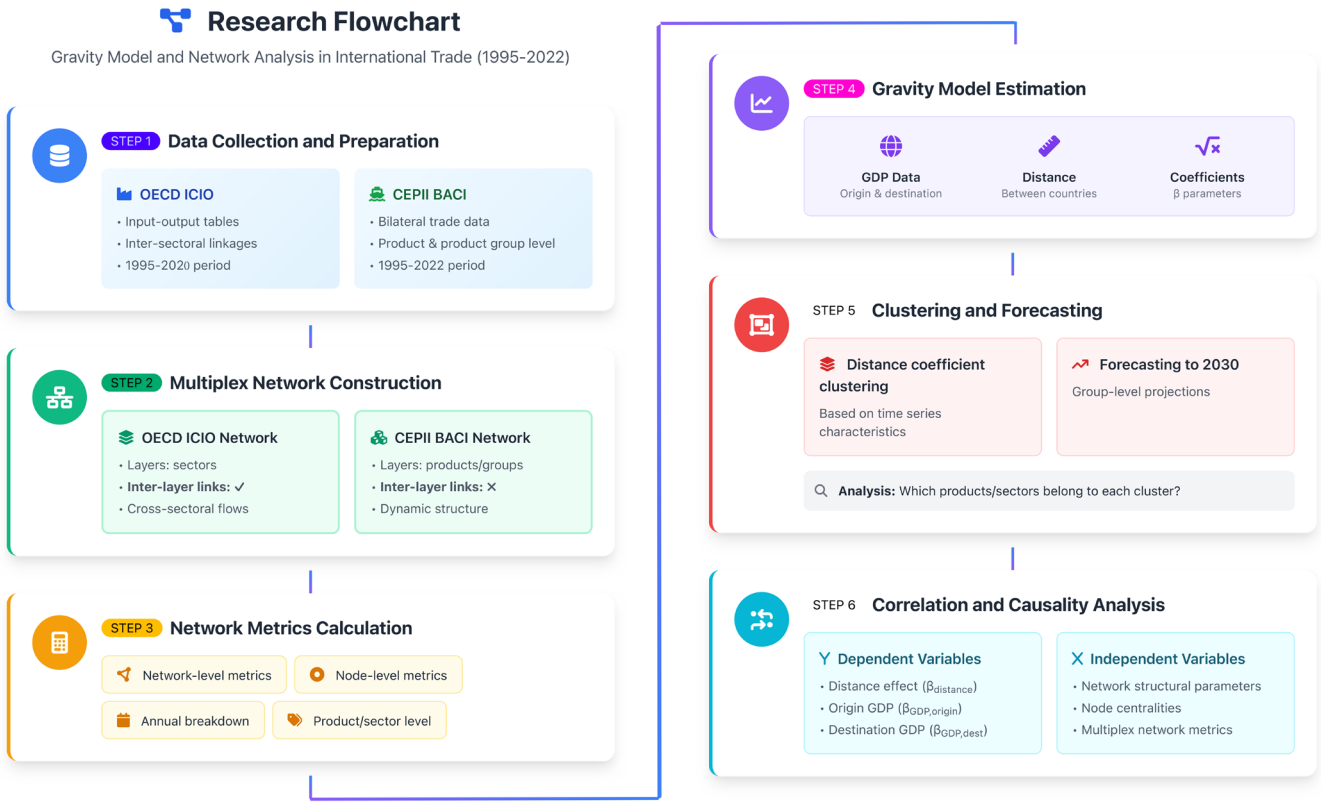


FIGURE 1 | Research flowchart. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 | Edge list structure for OECD-ICIO multiplex network.

Variable	Description
`Year`	Reference year $t \in \{1995, 1996, \ldots, 2022\}$
$\texttt{Node}_i$	Origin country (ISO 3-letter code)
$\texttt{Layer}_i$	Origin sector (ISIC Rev. 4 code)
$\texttt{Node}_j$	Destination country (ISO 3-letter code)
$\texttt{Layer}_j$	Destination sector (ISIC Rev. 4 code)
`Volume`	Transaction value in millions USD
`Quantity`	Physical quantity (where available)

TABLE 3 | Edge list structure for BACI-CEPII multiplex network.

Variable	Description
`Year`	Reference year $t \in \{1995, 1996, \ldots, 2022\}$
$\texttt{Product}_k$	Product identifier (HS-2, HS-4, or HS-6 code)
$\texttt{Node}_i$	Origin (exporting) country
$\texttt{Node}_j$	Destination (importing) country
`Volume`	Trade value in thousands USD
`Quantity`	Physical quantity in metric tons

The edge list structure is provided in Table 3.

Table 4 presents illustrative examples at different aggregation levels.

In these examples:

- HS-87: Vehicles other than railway or tramway rolling stock
- HS-8703: Motor cars and other motor vehicles for transport of persons
- HS-870323: Vehicles with spark-ignition engine, 1500–3000 cc
- HS-85: Electrical machinery and equipment
- HS-8517: Telephone sets and other apparatus for transmission/reception
- HS-851712: Telephones for cellular networks or wireless networks

TABLE 4 | Example entries: BACI-CEPII edge list at different aggregation levels.

Year	$\texttt{Product}_k$	$\texttt{Node}_i$	$\texttt{Node}_j$	Volume	Quantity
Product Group Level (HS-2)					
2020	87	DEU	USA	24,567,890	892,345
2020	85	CHN	DEU	45,678,123	1,234,567
Product Subgroup Level (HS-4)					
2020	8703	DEU	USA	18,234,567	654,321
2020	8517	CHN	DEU	12,345,678	456,789
Product Level (HS-6)					
2020	870,323	DEU	USA	8,765,432	234,567
2020	851,712	CHN	DEU	5,432,198	123,456

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4.1.3 | Constructed Analytical Databases

This section describes the analytical databases constructed from the multiplex networks, integrating network metrics with gravity model coefficients. This corresponds to **Step 3 (Network Metrics Calculation)**, **Step 4 (Gravity Model Estimation)**, and **Step 6 (Correlation and Causality Analysis)** of our research framework.

For each year t, product k (or sector s), we estimate the following gravity model:

$$Trade_{ij,t}^{k} = \exp\left(\alpha + \beta_{o,t}^{k} \ln\left(GDP_{i,t}\right) + \beta_{d,t}^{k} \ln\left(GDP_{j,t}\right) + \gamma_{t}^{k} \ln\left(Distance_{ij}\right) + \sum_{m=1}^{9} \delta_{m,t}^{k} D_{m,ij} \right) \epsilon_{ij,t} \quad (1)$$

We now estimate (1) by PPML (Santos Silva and Tenreyro 2006) rather than by log-linearized OLS: the dependent variable is trade in levels (not logged), which allows zero trade flows to enter the estimation without ad hoc adjustment, while the continuous regressors (distance and the two GDP terms) remain log-transformed so that $\beta_{o,t}^{k}$, $\beta_{d,t}^{k}$, and γ_{t}^{k} retain their interpretation as elasticities and stay directly comparable to the OLS estimates we report as a robustness check (Section Comparison of OLS and PPML Estimates for the Bilateral Control Variables). The term $\sum_{m=1}^{9} \delta_{m,t}^{k} D_{m,ij}$ adds nine time-invariant and time-varying bilateral control variables from the CEPII Gravity database (Conte et al. 2022): common official language (D_1, `comlang_off`), common ethnic language (D_2, `comlang_ethno`), contiguity (D_3, `contig`), a colonial relationship post-1945 (D_4, `col45`), an ever-colonial relationship (D_5, `col_dep_ever`), a common legal origin pre-transition (D_6, `comleg_pretrans`), joint free-trade-agreement/WTO membership (D_7, `fta_wto`), and EU membership of the origin (D_8, `eu_o`) and destination (D_9, `eu_d`) country. As we discuss in Section 8, standard country-pair or exporter/importer fixed effects cannot be added to (1) without changing our year-by-year cross-sectional design: with one observation per pair per year, a pair fixed effect would saturate the data, and an exporter- or importer-level fixed effect would be collinear with, and would absorb, $\ln\left(GDP_{i,t}\right)$ or $\ln\left(GDP_{j,t}\right)$—precisely the elasticities this study estimates. The nine CEPII covariates are therefore the most disaggregated bilateral controls that can be added without sacrificing identification of $\beta_{o,t}^{k}$, $\beta_{d,t}^{k}$, and γ_{t}^{k}.

The estimated coefficients that serve as dependent variables in the subsequent network analysis (Y) remain unchanged from the original design:

- $\beta_{o,t}^{k}$: Origin (exporter) GDP elasticity for product/sector k in year t
- $\beta_{d,t}^{k}$: Destination (importer) GDP elasticity for product/sector k in year t
- γ_{t}^{k}: Distance elasticity (distance dependence) for product/sector k in year t

The nine $\delta_{m,t}^{k}$ coefficients are estimated jointly with $\beta_{o,t}^{k}$, $\beta_{d,t}^{k}$, and γ_{t}^{k} in every cross-section, but—consistent with our focus on distance and market-size elasticities—are not themselves carried forward as Y variables into the network-causality analysis; we report their sign stability across years in Section Comparison of OLS and PPML Estimates for the Bilateral Control Variables instead of reproducing 9 additional coefficient panels in the main text.

Network-level and node-level indicators (X) computed for each year. Applied network-level indicators are listed in the Appendix A in Table A9 and node-level indicators are listed in Table A7. The two-dimensional panel database aggregates network-level metrics across all layers, providing a time-series perspective on global trade network evolution.

The two-dimensional panel database aggregates network-level metrics across all layers. Each row represents a year, with columns containing gravitational coefficients ($\beta_{o,t}$, $\beta_{d,t}$, γ_{t}) and network indicators (Density, Modularity, etc.). An example entry:

1995 | 0.842 | 0.756 | − 1.234 | 0.432 | 0.567 | …

where the values represent Year, $\beta_{o,t}$, $\beta_{d,t}$, γ_{t}, Density_{t}, and Modularity_{t} respectively.

Three-dimensional databases maintain product/sector disaggregation, enabling analysis of heterogeneity across trade categories. For the BACI-CEPII product-level database, each row represents a year-product combination with gravitational coefficients and node-level indicators:

1995 | 87 | 0.912 | 0.845 | − 1.456 | 45.2 | 0.234 | …

representing Year, Product (HS-2 code), $\beta_{o,t}^{k}$, $\beta_{d,t}^{k}$, γ_{t}^{k}, Degree Centrality, and Betweenness Centrality respectively.

Similarly, for the OECD-ICIO sector-level database:

1995 | C29 | 0.876 | 0.812 | − 1.534 | 38.4 | 42.1 | …

representing Year, Sector (ISIC code), $\beta_{o,t}^{s}$, $\beta_{d,t}^{s}$, γ_{t}^{s}, in-degree centrality, and out-degree centrality respectively.

Following **Step 6 (Correlation and Causality Analysis)**, we construct result matrices that capture the relationship between gravitational coefficients and network structural parameters across different time lags.

For each gravitational coefficient (distance, origin GDP, destination GDP), we create a two-dimensional matrix where *rows* represent network indicators (node-level and network-level metrics); *columns* represent sectors (OECD-ICIO) or product groups (BACI-CEPII) and *cell values* represent the optimal lag length (1–5 years) from Granger causality tests or a correlation value in a correlation analysis.

4.2 | Employed Network and Gravity Indicators

This section provides a consolidated overview of the indicators used throughout the empirical analysis, organized into dependent variables (gravity coefficients) and independent variables

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(network metrics). The indicator selection directly addresses 1 by incorporating not only conventional network measures (centrality, modularity, clustering) but also novel topological indicators—specifically, fractal dimensions and the four moments of local lacunarity distributions (mean, standard deviation, skewness, kurtosis)—that capture multiscale heterogeneity in trade network structure.

4.2.1 | Employed Network and Gravity Indicators

Following **Step 4** of the research framework (Figure 1), we estimate time-varying gravity coefficients for each year, sector, and product group. Table 5 summarizes these dependent variables.The calculations of gravity coefficients are provided in Section Employed Gravity Models.

Network indicators are computed following **Step 3** of the research framework. We employ both network-level metrics (characterizing the entire trade network) and node-level metrics (computed for individual countries and subsequently aggregated). Table 6 organises these indicators by the trade network phenomena they capture, while Table 7 explicates their theoretical connections to gravity model coefficients.The theoretical rationale for examining relationships between network phenomena and gravity coefficients stems from the interdependence between trade costs (captured by gravity parameters) and network topology (captured by structural indicators).

Understanding these directional relationships is essential for addressing our research question and for the policy implications discussed in C_3. For instance, if modularity increases (indicating regionalization), Granger-cause rising distance coefficients, this would suggest that trade bloc formation actively increases distance sensitivity rather than merely correlating with it.

The indicator groupings in Tables 6 and 7 provide the conceptual framework for interpreting the causality analysis results. By examining which phenomenon groups exhibit significant Granger-causal relationships with specific gravity coefficients, we can characterize the nature of network-gravity co-evolution and identify which structural changes precede or follow shifts in trade friction parameters.

TABLE 5 | Gravity model coefficients (dependent variables Y).

Symbol	Name	Interpretation
$\gamma_t^{(k)}$	Distance coefficient	Elasticity of trade with respect to geographic distance; captures trade friction intensity
$\beta_{o,t}^{(k)}$	Origin (exporter) GDP coefficient	Elasticity with respect to exporter economic mass; reflects supply-side capacity
$\beta_{d,t}^{(k)}$	Destination (importer) GDP coefficient	Elasticity with respect to importer economic mass; reflects demand-side market access

4.3 | Methodological Framework

This section outlines the integrated methodological approach employed to address our research question. Our framework combines multiple analytical techniques that mutually reinforce and validate findings through three forms of triangulation:

- **Data triangulation**: Two trade databases (OECD ICIO and CEPII-BACI) capturing different dimensions of international trade, directly supporting C_3;
- **Model triangulation**: Three causality testing approaches (traditional Granger, Bayesian, and Random Forest-based) with complementary strengths, implementing C_3;
- **Level triangulation**: Analyses at aggregate, sectoral, and product-group levels to document heterogeneity, enabling C_3.

This multi-method design addresses 1 by employing rigorous causality testing rather than correlation-based inference, which has been the dominant approach in prior literature (see Table 1). The methodological workflow follows Steps 4–6 of Figure 1.

4.3.1 | Gravity Model Estimation and Coefficient Extraction

We estimate parsimonious structural gravity models separately for each year t and each layer k (sector or product group) using the specification in Equation (A.2). This yields time series of coefficients $\left\{\gamma_t^{(k)}, \beta_{o,t}^{(k)}, \beta_{d,t}^{(k)}\right\}$ that serve as dependent variables in subsequent analyses. The cross-sectional estimation approach allows unrestricted temporal variation in elasticities, enabling detection of structural breaks and trend changes (Disdier and Head 2008; Yotov 2012). Detailed model specification is provided in Section Employed Gravity Models.

4.3.2 | Clustering and Forecasting

Following Step 5 of the research framework, we apply Generalized Network-based Dimensional Analysis (GNDA) (Kosztyán et al. 2024) to identify clusters of sectors and product groups exhibiting similar temporal patterns in their gravity coefficients. This clustering step is essential for addressing the heterogeneity dimension of our research question: rather than assuming all sectors or products follow identical network-gravity dynamics, we first identify groups with similar temporal trajectories and then examine whether these groups also share common causal relationships with network indicators.

The GNDA method offers two advantages for our analysis. First, it automatically determines the optimal number of clusters through modularity optimization, avoiding arbitrary specification. Second, by using Spearman correlation as the similarity function, it captures monotonic relationships in coefficient trajectories without requiring linearity assumptions.

Within each identified cluster, we construct representative time series using eigenvector centrality-weighted aggregations. This aggregation approach ensures that sectors or products most

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 6 | Network indicators grouped by trade network phenomena.

Phenomenon	Indicators	Interpretation of temporal changes
Market Concentration & Dominance	In-Degree Centralization (DZI), Out-Degree Centralization (DZO), Betweenness Centralization (BZ), PageRank Centralization (PRZ), Harmonic Centralization (HZ), Power Centralization (PZ); and mean values of Betweenness Centrality (BC), Eigenvector Centrality (EC), and PageRank Centrality (PRC)	Increasing centralization indicates growing dominance of hub economies in trade intermediation. Rising values signal higher systemic risk from hub dependence; declining values suggest more distributed, polycentric trade patterns
Trade Intensity & Connectivity	Density (Dens), AVPL, Global Efficiency (GLE), Average Local Efficiency (ALE); and mean values of In-degree Centrality (DCI), Out-degree Centrality (DCO), In-strength Centrality (SCI), and Out-strength Centrality (SCO)	Rising density and efficiency indicate deepening trade integration and reduced effective distance between economies. Increasing average path length signals fragmentation or emergence of barriers; declining connectivity reflects trade disintegration or regionalization
Clustering & Regionalization	Transitivity (Trans), Modularity Value (MOD), Assortativity (Assort)	Increasing modularity and transitivity indicate formation of regional trade blocs and preferential trading clusters. Rising assortativity suggests large economies increasingly trade with each other (rich-club formation); changes reflect shifting balance between globalization and regionalization
Hierarchical Structure & Complexity	Fractal Dimension (FDim), 1–4 moments of Local Lacunarity (LL) (LL_m1, LL_m2, LL_m3, LL_m4)	Fractal dimension captures self-similarity across scales; higher values indicate more complex, space-filling network structure. Lacunarity moments measure heterogeneity in trade partner distributions. Temporal changes signal structural phase transitions in network organization
Network Resilience & Vulnerability	RRes, SRes	Resilience (Random Attack) versus Resilience (Systematic Attack) reveals vulnerability to different disruption types. Declining resilience indicates increasing fragility; divergence between RRes and SRes signals hub-dependent structure vulnerable to targeted shocks
Local Trade Patterns	Motif (MOT), Trace; mean values of Hub Centrality (HC), Authority Centrality (AC), Katz Centrality (KC), and K-core (KCr)	Motif frequencies capture recurring local configurations (mutual, chain, cyclic patterns). Hub and authority centralities reveal directional influence in value chains. Changes indicate evolving micro-structure of trade relationships and value chain positioning

Note: Detailed definitions and calculation formulas are provided in Tables A9 and A7 of the Appendix A. Network-level indicators are listed first, followed by aggregated node-level indicators (prefixed with 'Avg.').

central to each cluster's characteristic pattern receive appropriate weight in subsequent causality analyses.

For forecasting gravity coefficients to 2030, we employ both traditional Autoregressive Integrated Moving Average (ARIMA) models and Bayesian ARIMA specifications. The forecasts serve two purposes related to C_3: they extend the temporal coverage of our coefficient series, and the Bayesian approach provides posterior predictive distributions that appropriately quantify forecast uncertainty given our relatively short time series (26–28 annual observations). These forecasts also enable forward-looking policy discussions regarding the expected evolution of distance sensitivity across different sector and product clusters. Technical details are provided in Section Forecastings.

The resulting clusters are labelled NDA_1, NDA_2, NDA_3 (and occasionally NDA_4 for finer product classifications), where the subscript indicates the cluster identifier based on modularity-optimized community detection. Sectors or products within the same NDA cluster exhibit similar temporal trajectories in their gravity coefficients. Cluster membership lists for all disaggregation levels are provided in the following link (https://github.com/kzst/GRAV_ICIO_CEPII).

4.3.3 | Correlation and Causality Analysis

The core analytical contribution addresses the central element of our research question: do changes in network structural properties Granger-cause shifts in gravity coefficients, or vice versa (**Step 6**)? This section describes three complementary causality testing approaches, each addressing different aspects of 1 regarding the absence of rigorous causality analysis in prior literature.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 7 | Theoretical links between network phenomena and gravity coefficients.

Phenomenon	Link to distance coefficient (γ)	Link to origin GDP (β_o)	Link to destination GDP (β_d)
Market Concentration & Dominance	High centralization may amplify or attenuate distance effects depending on hub locations; concentrated intermediation can compress effective distances	Dominant exporters (high hub centrality) may exhibit stronger GDP-trade elasticity due to economies of scale in trade infrastructure	Concentrated import markets (high authority) may show stronger demand-side elasticity from market access advantages
Trade Intensity & Connectivity	Higher density and efficiency indicate reduced friction; negative correlation with $\lvert\gamma\rvert$ expected as connectivity improvements compress distance effects	Increased connectivity raises trade capacity for given GDP; may strengthen or weaken β_o depending on whether capacity or competition effects dominate	Denser import networks provide more supplier options; may moderate β_d through diversification effects
Clustering & Regionalization	Higher modularity indicates regional bloc formation; may increase within-bloc distance insensitivity while raising between-bloc friction (heterogeneous γ)	Assortative mixing (large-with-large trade) may strengthen GDP elasticity within size classes while weakening cross-class relationships	Regional clustering affects market access differentially; destination GDP effects may vary by bloc membership
Hierarchical Structure & Complexity	Higher fractal dimension and lacunarity indicate multiscale heterogeneity; complex structure may generate non-linear distance effects or sector-specific patterns	Hierarchical structure implies differentiated roles; origin GDP effects may vary across hierarchy levels (core vs. periphery)	Complex destination structure suggests differentiated market access; β_d may capture hierarchy-dependent demand effects
Network Resilience & Vulnerability	Low resilience indicates fragile structure; distance sensitivity may increase as redundant pathways disappear (fewer alternatives to direct routes)	Vulnerable networks may show weaker GDP elasticity as supply disruption risks discourage trade expansion	Fragile import networks may exhibit stronger β_d as importers concentrate on secure, large-market sources
Local Trade Patterns	Motif structures (chains, cycles) indicate value chain organization; distance effects may operate through intermediate nodes rather than direct bilateral links	Hub centrality reflects upstream positioning; high hub scores may strengthen origin GDP effects through supply chain multipliers	Authority centrality reflects downstream positioning; high authority may amplify destination GDP effects through derived demand channels

Note: These theoretical connections motivate the empirical causality analysis. Further elaboration on indicator-coefficient connections is provided in Table A10 of the Appendix A.

4. *Traditional Granger Causality*: We implement classical nested F-tests comparing restricted (autoregressive) and unrestricted (including lagged network indicators) models (Granger 1969). For panel data with multiple sectors or products, we extend this to the Dumitrescu and Hurlin (2012) framework incorporating unit fixed effects. A stringent significance level ($\alpha = 0.01$) controls for multiple testing across numerous variable-lag combinations. This approach directly tests whether past values of network indicators (X_{t-k}) help predict current gravity coefficients (Y_t) beyond what is predictable from the coefficients' own past values. A finding that network indicator X Granger-causes gravity coefficient Y at lag k implies that structural changes in trade networks temporally precede—and potentially drive—shifts in trade frictions or market access elasticities. Technical details are provided in Section Traditional and Bayesian Granger Causality Analysis.

5. *Bayesian Granger Causality*: Bayesian model comparison via Bayes factors (Kass and Raftery 1995) complements the frequentist approach by providing continuous evidence measures rather than binary reject/fail-to-reject decisions. Using Jeffreys–Zellner–Siow priors (Liang et al. 2008), we compute marginal likelihoods for restricted and unrestricted models, requiring BF ≥ 5 for substantial evidence of causality. The Bayesian approach offers three advantages for our analysis. First, it naturally handles parameter uncertainty, which is substantial given our short time series. Second, it avoids the multiple testing issues inherent in sequential hypothesis testing across many indicator-coefficient pairs. Third, the continuous Bayes factor scale enables nuanced interpretation of evidence strength rather than arbitrary significance thresholds. Technical details are provided in Section Traditional and Bayesian Granger Causality Analysis.

6. *Random Forest-Based Causality*: To capture potential non-linear relationships missed by linear specifications, we employ Random Forest regression with a Boruta-inspired shadow feature approach (Breiman 2001; Kursa and Rudnicki 2010). Variable importance is assessed via permutation importance, with causality declared when the

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

importance of lagged network indicators exceeds that of randomly permuted shadow features. This nonparametric method accommodates interactions and nonlinearities without prespecification, addressing 1's emphasis on moving beyond correlation-based methods. If network-gravity relationships involve threshold effects or complex interactions, Random Forest-based tests will detect patterns that linear Granger tests miss. Technical details are provided in Section Random Forest-Based Granger-Like Causality Analysis.

The three approaches constitute methodological triangulation central to C_3. Relationships detected consistently across all methods provide robust evidence of genuine causal linkages. Method-specific findings reveal the nature of relationships: results appearing only in traditional and Bayesian tests suggest linear dynamics, while Random Forest-exclusive findings indicate nonlinear mechanisms.

For panel analyses at sectoral and product-group levels, we construct result matrices where rows represent network indicators, columns represent economic categories, and cell values indicate the optimal lag at which Granger causality is detected. These matrices directly enable C_3 by revealing systematic heterogeneity: which sectors or products show network-leading versus gravity-leading dynamics, and at what temporal horizons. The technical specifications for panel causality tests are detailed in Section Panel Version of Granger-Causality Tests.

Table 8 summarizes how each methodological component addresses specific research gaps and contributes to answering our research question.

This integrated framework ensures that each analytical step contributes to answering specific aspects of our research question while addressing identified gaps in the literature. Follow this methodological structure, first presenting temporal patterns and forecasts (addressing the 'how do coefficients evolve' dimension), then correlation analyses (establishing contemporaneous associations), and finally causality tests (determining temporal precedence between network structure and gravity parameters).

5 | Results

5.1 | Temporal Patterns and Forecastings of Gravity Coefficients

Figure 2 presents the time series of gravity coefficients estimated from the aggregate OECD-ICIO trade network for 1995–2022, with ARIMA-based forecasts extending to 2030.

Figure 2a shows that the distance coefficient (γ_t) remained relatively stable, fluctuating between approximately − 0.66 and − 0.63 over 1995–2020 (from − 0.66 in 1995 to − 0.64 in 2020), with the point forecast holding essentially flat at about − 0.64 through 2030—indicating a modest, and by 2020 largely completed, decline in distance sensitivity rather than a continuing trend. Figure 2b presents the origin ($\beta_{o,t}$) and destination ($\beta_{d,t}$) GDP coefficients, which range from approximately 0.47 to 0.58 and generally increase over time—from 0.50 to 0.55 for the origin coefficient and from 0.51 to 0.57 for the destination coefficient between 1995 and 2020—with the origin coefficient somewhat more volatile, particularly in the mid-2000s. Forecast intervals widen after 2020, reflecting greater uncertainty in out-of-sample predictions, even though the point forecasts themselves remain flat. The sectorial forecasts and the forecasts of product groups are presented in Section Forecasting of Elasticies. Out-of-sample forecast accuracy for all series is reported in Table A1.

5.2 | Forecastings of Network Indicators

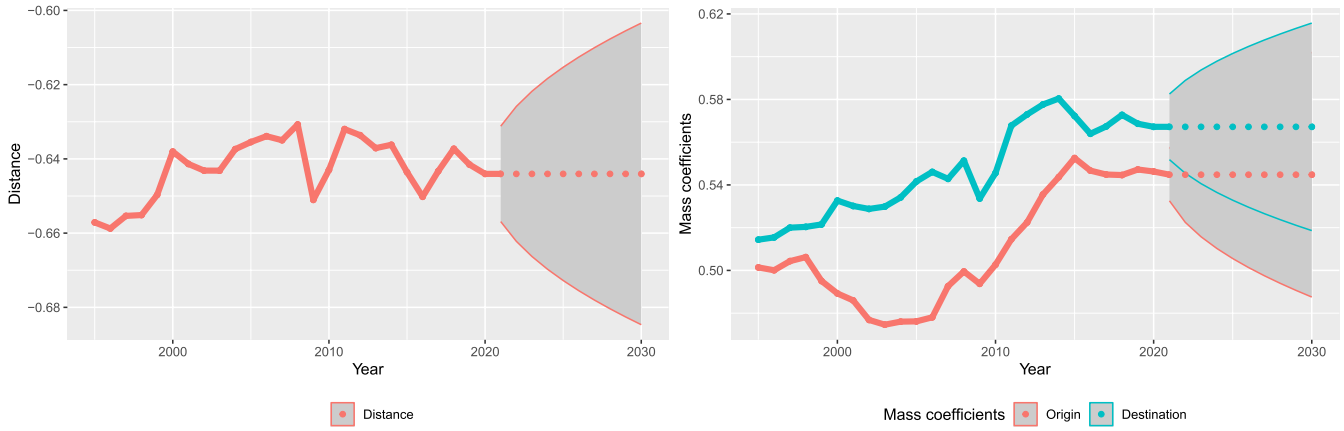
Figure 3 presents the temporal evolution of key network-level indicators computed from the OECD-ICIO multiplex trade network, with ARIMA-based forecasts to 2030.

Figure 3a displays centralization measures and modularity: in-degree (DZI) and out-degree (DZO) centralization remain relatively stable around 0.1–0.2, while betweenness centralization (BZ) shows higher values near 0.4, indicating persistent hub-and-spoke structures. Modularity (MOD) fluctuates between 0.7 and 0.9, suggesting strong community structure throughout

TABLE 8 | Mapping methodological components to research objectives.

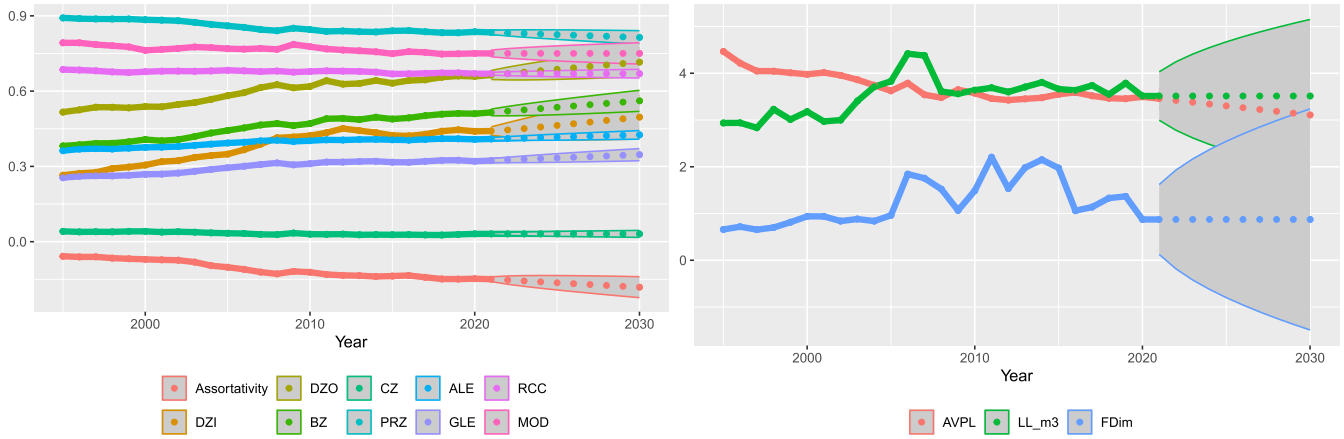
Method	Research gap addressed	Contribution	RQ dimension
Dual database design (OECD-ICIO, BACI-CEPII)	RG_3: Cross-database validation	C_1: Time series at multiple aggregation levels	Temporal evolution
Fractal dimensions & lacunarity moments	RG_2: Novel network indicators	C_2: Comprehensive indicator set	Network structure
Traditional Granger causality	RG_1: Rigorous causality testing	C_2: Directional relationships	Causal precedence
Bayesian Granger causality	RG_1: Handling short panels	C_2: Uncertainty quantification	Causal precedence
Random Forest causality	RG_1: Nonlinear relationships	C_2: Model triangulation	Causal precedence
Panel causality by sector/product	RG_4: Structural heterogeneity	Network Resilience and Shock Propagation: Systematic patterns	Heterogeneity
GNDA clustering	—	C_1: Temporal pattern groups	Heterogeneity
Bayesian ARIMA forecasting	—	C_1: Extended time coverage	Temporal evolution

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

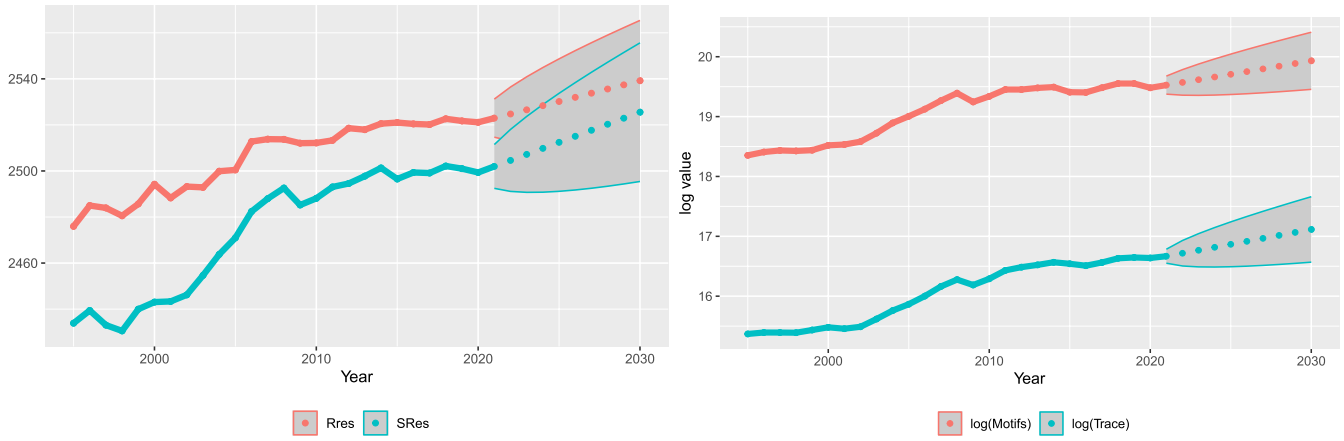


(a) Distance elasticity (b) GDP elasticities

FIGURE 2 | Aggregate gravity coefficients from OECD-ICIO network (1995–2022) with ARIMA forecasts to 2030 (95% CI). [Colour figure can be viewed at wileyonlinelibrary.com]



(a) Centralization indices and modularity (b) AVPL, fractal dimension, and lacunarity



(c) Network resilience (random vs. targeted attack (d) Motif count and trace (log scale)

FIGURE 3 | Network-level properties (forecast with ARIMA, 95% confidence interval, based on OECD-ICIO database). [Colour figure can be viewed at wileyonlinelibrary.com]

the period. Figure 3b shows that average path length (AVPL) remains stable around 2, fractal dimension (FDim) varies between 3 and 4, and the third lacunarity moment (LL_m3) exhibits considerable volatility. Figure 3c presents resilience measures: both Resilience (Random Attack) (RRes) and Resilience (Systematic Attack) (SRes) fluctuate between 2460 and 2540, with SRes consistently slightly lower, indicating moderate vulnerability to targeted disruptions. Figure 3d shows that both motif counts and trace (self-loops representing domestic transactions) increase steadily over time on a logarithmic scale, from e^{16} to e^{20}, reflecting the growing complexity and volume of trade relationships.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

5.3 | Correlation Analysis

Figure 4 presents Spearman correlation coefficients between the three gravity coefficients (origin GDP, destination GDP, and distance) and 29 network-level indicators, computed from the aggregate OECD-ICIO time series.

Strong positive correlations (0.8–1.0) are observed between both GDP coefficients and centralization measures (DZI, DZO, BZ, CZ, PRZ, PZ), density, efficiency measures (ALE, GLE), and transitivity, while the distance coefficient shows strong negative correlations (− 0.7 to − 0.8) with these same indicators. The lacunarity moments (LL_m1 through LL_m4) exhibit moderate positive correlations (0.6–0.8) with GDP coefficients and negative correlations with distance. Notably, modularity (MOD) shows weak correlations with all gravity coefficients (around − 0.1 to 0.3), while fractal dimension (FDim) displays moderate negative correlations (− 0.4 to − 0.5) with GDP coefficients. Similarity indicators (NS, ES, PD, SD, SP) show strong positive correlations with GDP coefficients, suggesting that structural stability in the trade network coincides with higher economic mass elasticities.

The sector-level Spearman correlations at the full, non-aggregated resolution are reported in Appendix A: Figure A5; we discuss here the NDA-cluster-aggregated summary (Figure 5), which conveys the same pattern more legibly.

Figure 5 summarizes correlations between gravity coefficients and modularity, lacunarity moments, and fractal dimension for each identified sectoral cluster.

For distance coefficients (Figure 5a), all three sectoral clusters show a similarly moderate negative correlation with modularity (− 0.29 to − 0.34) and a weak positive correlation with fractal dimension (0.06 to 0.23); the clusters are therefore not clearly differentiated on these two indicators, in contrast to the more heterogeneous pattern found for the individual lacunarity moments, where LL_m1 ranges from − 0.11 for NDA_3 to 0.26 for NDA_1. For origin coefficients (Figure 5b), NDA_2 shows the strongest modularity correlation (0.23), somewhat above NDA_1 (0.14), while NDA_3 is the only cluster with a negative modularity correlation (− 0.21); the first two lacunarity moments are negative for NDA_1 and NDA_2 (− 0.10 to − 0.28) but positive for NDA_3 (0.04 to 0.08). Destination coefficients (Figure 5c) are not uniformly weak: although most of the 18 cluster–indicator correlations remain below 0.2 in absolute value, several exceptions reach magnitudes comparable to the strongest distance- or origin-side correlations—NDA_3's modularity correlation (− 0.41), NDA_2's fractal-dimension correlation (0.31), and NDA_1's third and fourth lacunarity moments (0.25 and 0.29)—so destination-side network sensitivity should not be characterized as uniformly weaker than the distance or origin side.

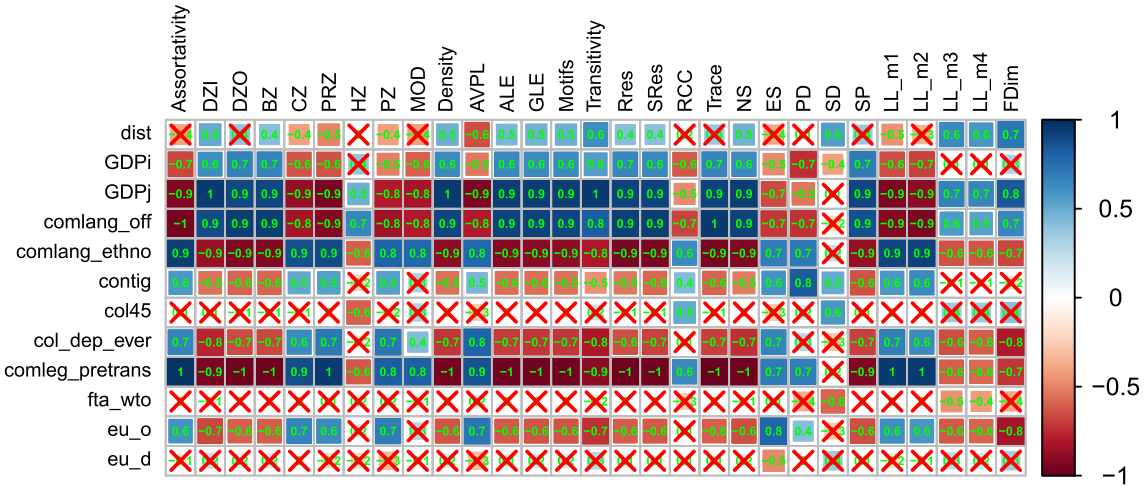


FIGURE 4 | Spearman's correlation between gravity coefficients and network-level (and aggregated node-level) indicators (based on OECD-ICIO database); cells marked with a red × are not statistically significant after Holm adjustment ($p_{adj} \geq 0.05$). [Colour figure can be viewed at wileyonlinelibrary.com]



FIGURE 5 | Spearman's correlation between gravity coefficients of clustered industries and modularity, lacunarity, and fractal dimension indicators. [Colour figure can be viewed at wileyonlinelibrary.com]

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The product-group-level Spearman correlations at the full, 86-product resolution are reported in Appendix, Figure A6; we discuss here the NDA-cluster-aggregated summary (Figure 6), which is more legible at this scale.

Figure 6 summarizes correlations between gravity coefficients and key network indicators for each identified product group cluster.

For distance coefficients (Figure 6a), the three product-group clusters are clearly differentiated: NDA_1 shows a consistent, moderate-to-strong pattern across indicators—negative correlations with all four lacunarity moments (− 0.34 to − 0.41), positive correlations with resilience measures (RRes, SRes around 0.41), and a positive correlation with GLE (0.39)—whereas NDA_2 and NDA_3 show only weak correlations with all of these indicators (magnitudes below 0.10 for the lacunarity moments and below 0.09 for the resilience measures). For origin coefficients (Figure 6b), correlations are broadly similar in magnitude across all three clusters rather than concentrated in one of them: NDA_2 and NDA_3 show the strongest GLE correlations (0.42 each) and the strongest lacunarity-moment correlations (− 0.42 to − 0.46 for LL_m1/LL_m2), while NDA_1 is comparable but somewhat weaker on these same indicators (GLE = 0.24, LL_m1 = − 0.28); all three clusters show a moderate positive correlation with resilience measures (0.35–0.50). For destination coefficients (Figure 6c), all three clusters are again identified (not two): NDA_1 and NDA_2 show a similar pattern of moderate positive correlation with GLE (0.35–0.40) and resilience measures (0.30–0.37) together with moderate-to-strong negative correlations with the third and fourth lacunarity moments (− 0.39 to − 0.45), whereas NDA_3 shows near-null correlations with most indicators except for a moderate negative correlation with assortativity (− 0.26).

5.4 | Causality Analysis

The causality results allow a direct assessment of the three hypotheses formulated in the Theoretical Framework (Section 2). Because our year-specific gravity coefficients now come from the PPML specification with the nine CEPII bilateral controls (Section 4.1.3), we re-estimated all three causality tests reported below on the revised coefficient series.

- *H_1 (Centralization → GDP elasticities)*: Supported. PageRank centralization (PRZ), density, and out-degree centralization (DZO) show significant Granger-causal relationships with origin GDP elasticity across the traditional and Random-Forest methods (Figure 7c), at a one-year lag in the traditional test and at four-to-five-year lags once we additionally require the Random-Forest shadow-feature significance threshold to be cleared (Section Random Forest-Based Granger-Like Causality Analysis). Of the 29 network indicators considered, 23 show a statistically significant traditional Granger relationship with origin GDP elasticity, indicating that this is a broad, not an isolated, pattern.
- *H_2 (Community structure → distance elasticity)*: Only weakly supported at the aggregate level. Of the 29 network indicators, only Average Path Length (AVPL, lag 5), Power Centralization (PZ, lag 4), Average Local Efficiency (ALE, lag 3), and Trace (lag 5) show a statistically significant traditional Granger relationship with the aggregate distance elasticity; modularity (MOD) itself is not among them, and none of these four indicators clears the Random-Forest shadow-feature threshold either (Section Random Forest-Based Granger-Like Causality Analysis). This aggregate-level null result is consistent with—rather than contradicting—our most robust finding: the panel causality results disaggregated by sector (Figure 9) and by product group (Figure 10) show that network structure does Granger-precede distance elasticity for specific groups of sectors and products, but that this relationship is heterogeneous enough across groups that it washes out when all 45 sectors (or 97 product groups) are pooled into a single aggregate network series.
- *H_3 (Network heterogeneity → GDP elasticities)*: Supported for GDP elasticities, not for distance elasticity, at the aggregate level. Local lacunarity moments (LL_m1–LL_m3) and fractal dimension (FDim) are among the statistically significant traditional Granger-causal predictors of both origin and destination GDP elasticity (Figure 7c,d), each at a one-year lag. None of the four lacunarity moments or FDim, however, shows a significant traditional Granger relationship with the aggregate distance elasticity, and the same holds after applying the Random-Forest shadow-feature correction (Section Random Forest-Based Granger-Like Causality Analysis). As with H_2, this suggests that network heterogeneity's link to distance elasticity, where it exists, operates at the sectoral and product level (Section 6.5.3) rather than at the aggregate network level.

For the distance coefficient (Figure 7), 17 of 29 network indicators show significant Granger-causal relationships. The strongest predictors are LL_m3 (7.1% relative importance, lag 4), Trace (6.9%, lag 3), and PageRank centralization PRZ (6.8%, lag 2). Lacunarity moments and path-related measures dominate, while centralization measures show mixed results. Figure 7b displays the Granger-causal network among explanatory variables, revealing interconnected relationships particularly among

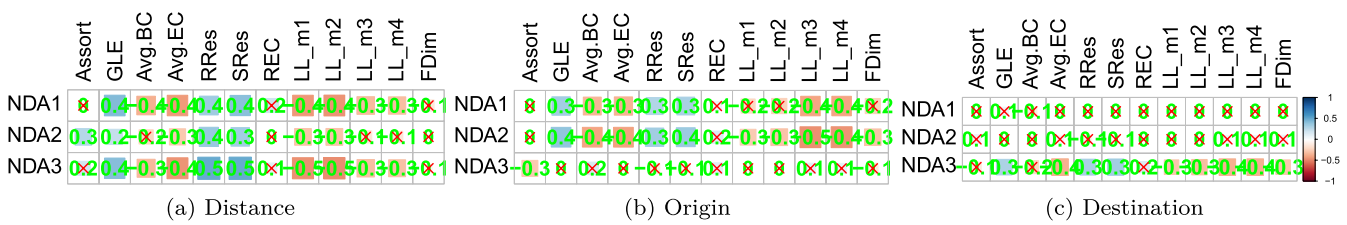
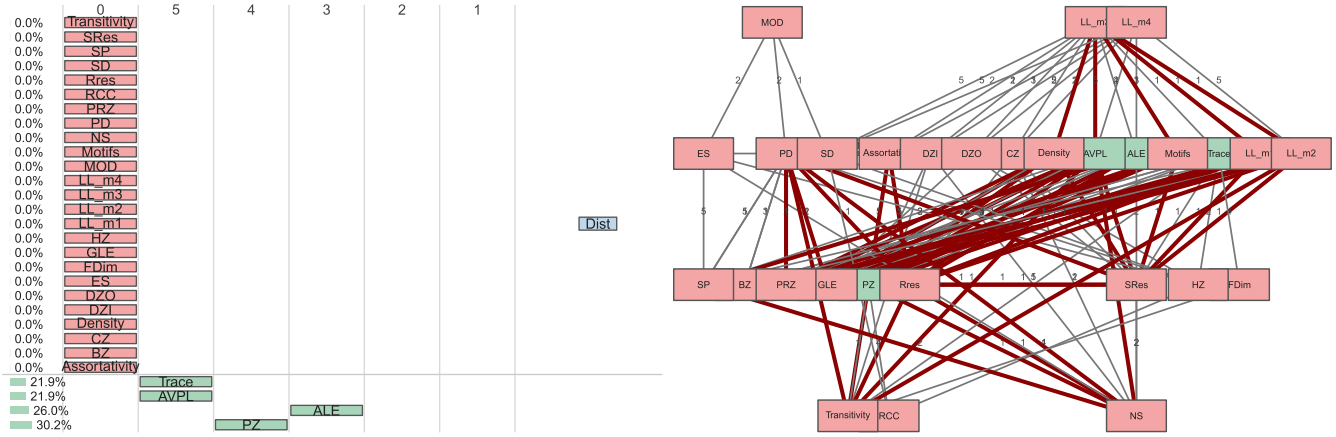


FIGURE 6 | Significant spearman's correlation values between gravity coefficients of clustered product groups and network-level/aggregated node-level indicators. [Colour figure can be viewed at wileyonlinelibrary.com]

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

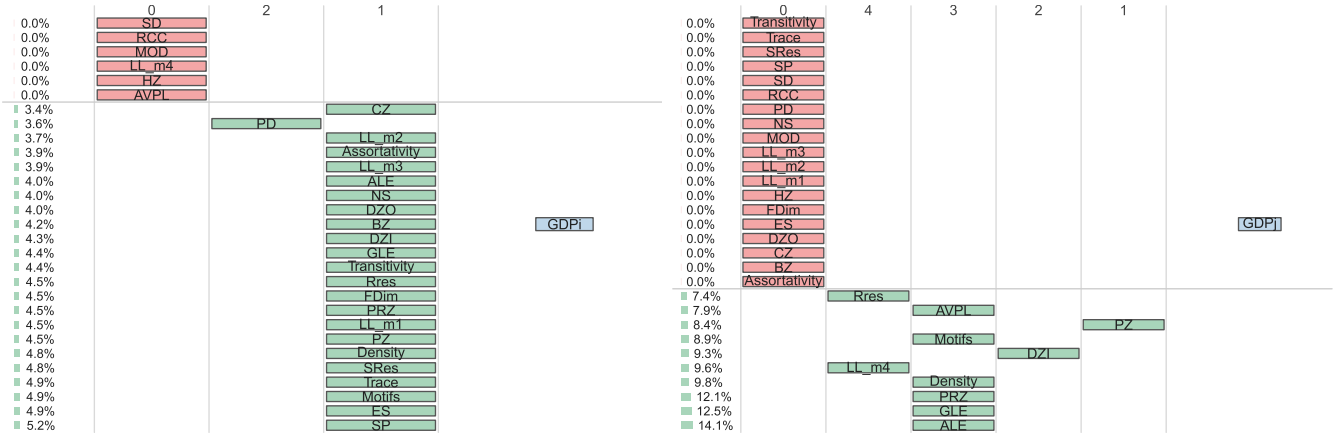
centralization measures. For origin GDP (Figure 7c), only five indicators achieve significance: ALE (23.1%), MOD (21.4%), NS (19.8%), PZ (18.8%), and AVPL (16.9%), all at lags 2–4. For destination GDP (Figure 7d), the same five indicators are significant with similar importance weights, suggesting that origin and destination GDP coefficients respond to similar network structural changes but with potentially different lag structures.

Figure 8 presents the Random Forest-based causality analysis results, which capture potential nonlinear relationships between network indicators and gravity coefficients that traditional linear Granger tests may miss.



(a) Granger-causal predictors of distance elasticity

(b) Granger-causal network among predictors

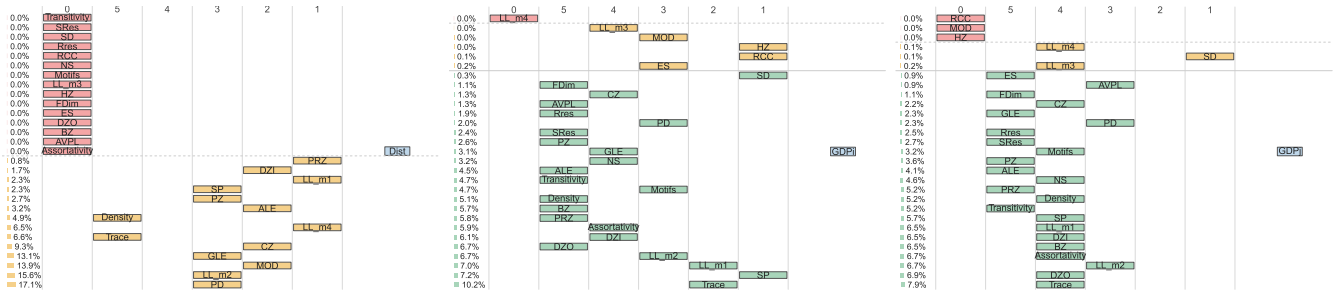


(c) Granger-causal predictors of origin GDP elasticity

(d) Granger-causal predictors of destination GDP elasticity

FIGURE 7 | Traditional Granger causality: From network indicators to gravity coefficients (OECD-ICIO aggregate). [Colour figure can be viewed at wileyonlinelibrary.com]

Figure 8 presents the Random Forest-based causality analysis results, which capture potential nonlinear relationships between network indicators and gravity coefficients that traditional linear Granger tests may miss. Because raw Random Forest importance can appear positive for an uninformative predictor purely by chance in a 26–28 observation panel, we report results after applying the shadow-feature significance



(a) Distance

(b) Origin

(c) Destination

FIGURE 8 | Random Forest-Based Granger-Like Causality Analysis between gravity and multilayer network indicators (based on OECD-ICIO database). [Colour figure can be viewed at wileyonlinelibrary.com]

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

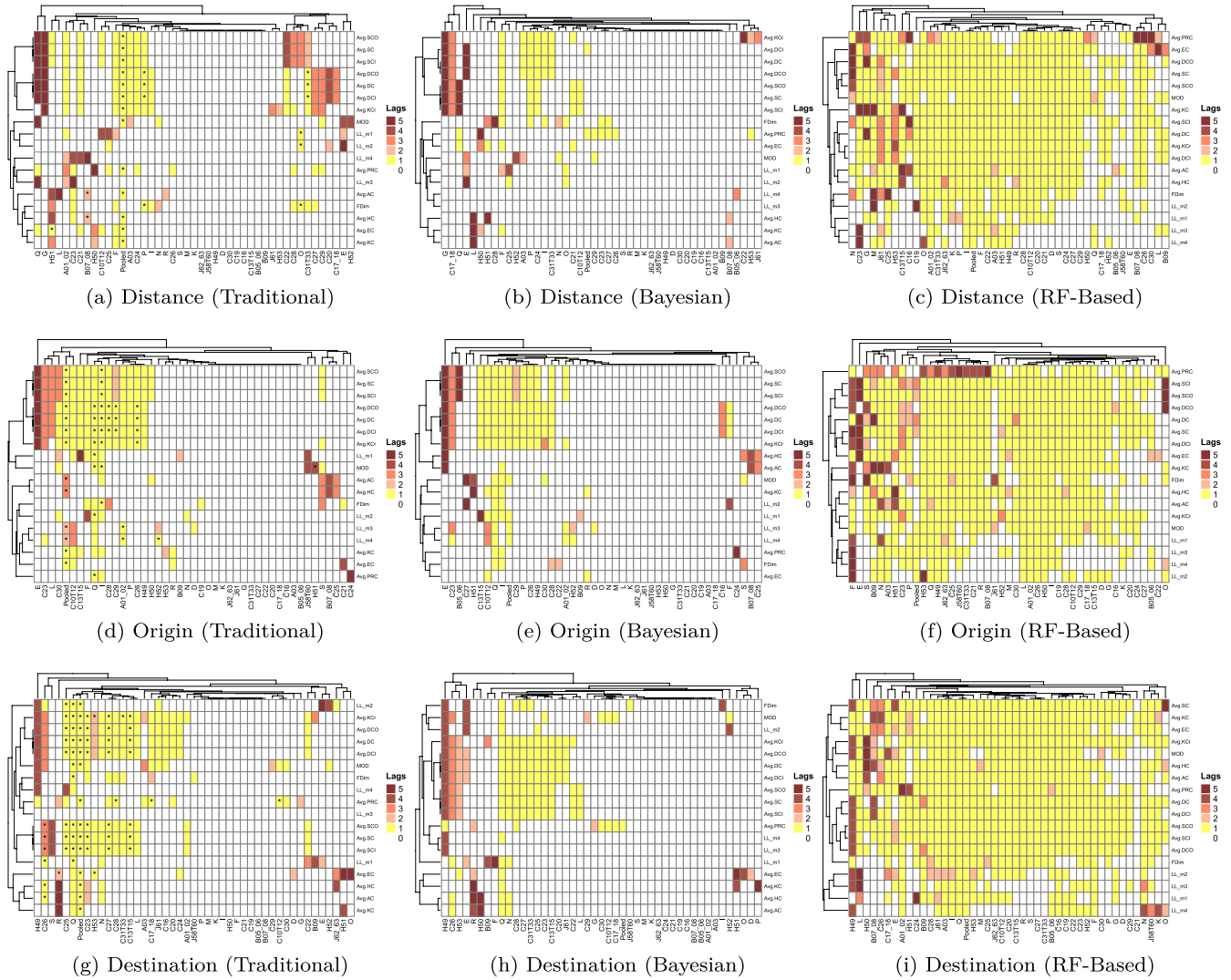


FIGURE 9 | Panel Granger causality: From network indicators to sectoral gravity coefficients (OECD-ICIO). [Colour figure can be viewed at wileyonlinelibrary.com]

threshold described in Section Random Forest-Based Granger-Like Causality Analysis (Equation A.57), rather than the raw importance ranking alone.

After shadow-feature correction, none of the 29 network indicators shows a statistically significant Random-Forest-based relationship with the aggregate distance coefficient (0 of 29; 14 of 29 had shown positive raw importance before correction). For origin GDP elasticity (Figure 8b), 23 of 29 indicators remain significant after correction, led by Trace, Shortest-Path similarity (SP), and the first two lacunarity moments (LL_m1, LL_m2) by total variable importance. For destination GDP elasticity (Figure 8c), 23 of 29 indicators likewise remain significant, again led by Trace, together with out-degree centralization (DZO), LL_m2, and Assortativity. Modularity (MOD), Harmonic closeness centralization (HZ), the Rich-Club Coefficient (RCC), and the fourth lacunarity moment (LL_m4) are the only indicators that fail to reach significance for *both* GDP directions, a consistent negative result across methods. As with the traditional Granger results (Section 5.4), the shadow-corrected Random Forest results show the same asymmetry: network structure predicts GDP elasticities broadly and robustly, but predicts the aggregate distance elasticity only weakly, if at all—reinforcing our decision to test the distance-elasticity channel primarily at the sectoral and product-group level (Section 6.5.3). The full shadow-correction decision table is in Appendix A: Table A2.

Figure 9 presents panel Granger causality results examining which network indicators predict changes in gravity coefficients across the 45 industrial sectors, using three complementary methods: traditional F-tests, Bayesian model comparison, and Random Forest importance. Cell colours indicate the lag (1–5 years) at which significant causality is detected; white cells indicate no significant relationship.

For distance coefficients (Figure 9a–c), substantial heterogeneity emerges across sectors. Across all 44 sectors, the Random Forest approach identifies a significant relationship in 588 of 810 sector-by-indicator cells for distance elasticity (72.6%), compared with 175 (21.6%) for the traditional test and 107 (13.2%) for the Bayesian test; the same pattern holds for origin and destination GDP elasticity (75.1% and 79.1% for Random Forest vs. 19%–25% for the linear methods). This three-to-four-fold increase in detected relationships is consistent with our aggregate-level finding

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

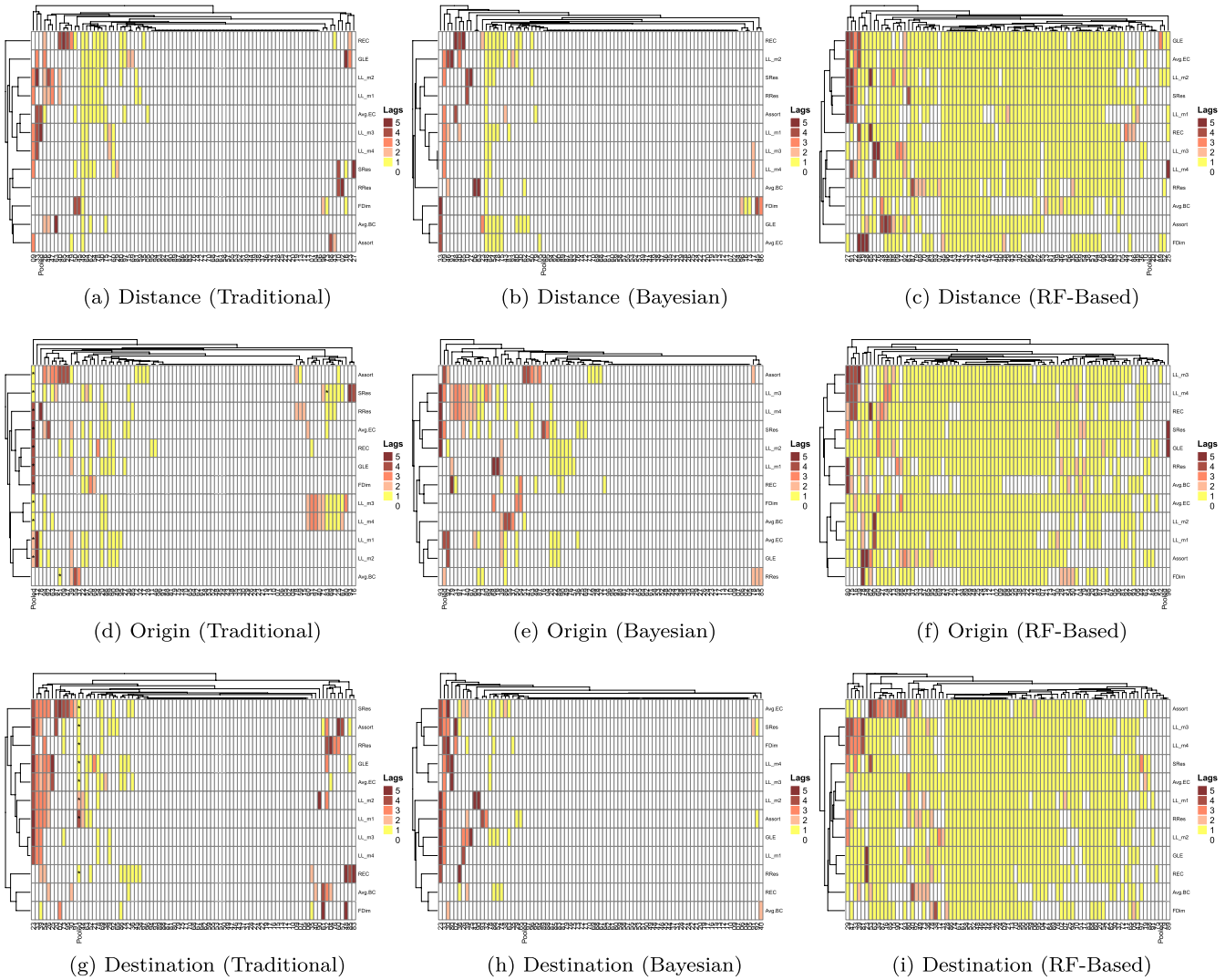


FIGURE 10 | Panel-based causality analysis for product groups (based on BACI-CEPII database). [Colour figure can be viewed at wileyonlinelibrary.com]

(Section A.12.3) that Random-Forest-based tests recover substantially more structure than linear Granger causality and supports treating the sectoral heterogeneity documented here as evidence of predominantly nonlinear, sector-specific network-to-gravity linkages rather than a uniform linear mechanism.

Figure 10 presents panel Granger causality results for the 97 product groups from the CEPII-BACI database, using the same three methodological approaches as the sectoral analysis.

For distance coefficients (Figure 10a–c), the traditional approach (Figure 10a) identifies SRes, LL_m2, LL_m1, and REC as significant predictors for the pooled panel, with considerable variation across individual product groups. Primary products (01–27) show stronger causality from resilience measures, while manufactured goods (84–90) respond more to centrality indicators. The Bayesian approach (Figure 10b) is more conservative, confirming LL_m2, LL_m1, and GLE as robust predictors. The RF-based approach (Figure 10c) provides a more comprehensive picture, identifying LL_m1, Reciprocity (REC), SRes, and RRes as top predictors with relatively even importance across product groups.

For origin GDP coefficients (Figure 10d–f), reciprocity and lacunarity moments emerge as consistent predictors across methods. The pooled results show that FDim, Avg.EC, and LL measures achieve significance, with agricultural products (01–24) showing stronger responses to structural changes than industrial products.

For destination GDP coefficients (Figure 10g–i), GLE and Avg.EC dominate the traditional and Bayesian results, while the RF approach identifies SRes, Avg.EC, and LL_m4 as important. Consumer goods (61–67, 94–96) exhibit stronger causality patterns than intermediate products, suggesting that final goods trade is more sensitive to network structural changes affecting market access.

6 | Discussion

This section interprets the empirical findings in light of the research question and contributions, organizing the discussion around four interconnected phenomena: (i) the temporal evolution of gravity coefficients, (ii) heterogeneity across sectors and

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

products, (iii) the relationship between network structure and gravity parameters, and (iv) the nature of causal linkages revealed by different methodological approaches.

6.1 | Temporal Dynamics of Distance and Mass Elasticities

Our analysis reveals systematic temporal patterns in gravity coefficients that address the first component of the research question regarding how income and distance elasticities evolve over 1995–2022.

At the aggregate network level, Figure 2 demonstrates that distance coefficients have exhibited a small decline in absolute magnitude, from approximately − 0.66 in the mid-1990s to approximately − 0.64 by 2020, with the point forecast flattening out rather than continuing to decline through 2030 (the widening confidence band reflects growing forecast uncertainty, not a projected further fall in distance sensitivity). This finding provides, at best, weak support for the hypothesis that 'the world is getting smaller'—the aggregate decline is modest, appears to have already levelled off by the end of the observation window, and, as the sectoral and product-level results below show, is far from universal. The mass coefficients (origin and destination GDP elasticities) shown in Figure 2b display a modest upward trend, from about 0.50–0.51 in the mid-1990s to about 0.55–0.57 by 2020, rather than being flat; economic mass therefore appears to have become, if anything, a somewhat *stronger* determinant of trade flows over the sample period even as distance friction weakened slightly.

These aggregate patterns mask substantial heterogeneity that becomes apparent at disaggregated levels. The sectoral analysis (Figure A1) reveals three distinct clusters with divergent temporal trajectories for distance elasticity—grouped by similarity in the *trajectory* of $\gamma_t^{(k)}$ over time rather than by a priori industry classification, which is why each cluster mixes sectors that might not share an obvious common label. The NDA_1 cluster includes telecommunications (J61) and IT services (J62_63) alongside several manufacturing and utility sectors (e.g., C23, C30, E, G) and exhibits the steepest decline in distance sensitivity; the presence of ICT-enabled services in this cluster remains consistent with theoretical predictions that such services should be least constrained by geographic distance (Lendle et al. 2016), though the cluster is not a pure 'services' group. The NDA_3 cluster—which now groups financial activities (K) together with core heavy-manufacturing sectors (C24–C29) and construction (D)—maintains relatively stable and higher distance elasticities, consistent with the continued importance of transport costs for physically bulky goods (Hummels and Klenow 2005) for its manufacturing members, though the inclusion of K suggests the shared pattern may instead (or additionally) reflect a common transaction-cost or regulatory-intensity channel that we do not separately test here.

This heterogeneity extends to product groups (Figure A2), where we identify three clusters with heterogeneous and, in NDA_1, steeply declining distance elasticities. These findings extend the literature on product-level gravity heterogeneity (Berthelon and Freund 2008; Baldwin and Harrigan 2011) by documenting not just cross-sectional differences but distinct temporal trajectories that may reflect differential impacts of technological change and trade liberalization across product categories.

The origin and destination GDP elasticities (Figures A1b,c and A2b,c) also exhibit cluster-specific patterns. Notably, sectors in NDA_1 for origin coefficients—including many knowledge-intensive services—show increasing GDP elasticity over time, suggesting that larger economies have become increasingly dominant as exporters in these sectors. This pattern aligns with theories of increasing returns and agglomeration economies in service trade (Helpman et al. 2008).

6.2 | Network Structure Co-Evolution

The temporal evolution of network-level properties (Figure 3) reveals structural changes that coincide with shifts in gravity coefficients, addressing the second part of our research question regarding the relationship between network topology and gravity parameters.

Figure 3a shows that centralization measures (DZI, DZO, BZ, PRZ) have generally increased over the observation period, indicating growing concentration of trade flows through hub economies. Simultaneously, modularity (MOD) has risen from approximately 0.4 in 1995 to over 0.6 by 2022, suggesting strengthening regional trade bloc structures. These patterns are consistent with the 'regionalization within globalization' thesis (Baldwin and Taglioni 2006) and provide network-structural evidence for the simultaneous operation of globalizing forces (declining distance sensitivity in aggregate) and regionalizing forces (increasing modularity).

The fractal dimension (FDim) and lacunarity measures (Figure 3b) offer novel insights not previously documented in the trade network literature. The fractal dimension has remained relatively stable around 2.5–3.0, suggesting that the trade network maintains self-similar properties across scales despite substantial reorganization. However, the third moment of lacunarity (LL_m3) has exhibited notable variation, with periods of increasing heterogeneity in local network structure coinciding with major trade disruptions (the 2008–2009 financial crisis and the post-2018 trade tensions).

Network resilience measures (Figure 3c) reveal an important asymmetry: RRes has remained stable while SRes has shown greater volatility. This divergence indicates that while the trade network maintains robustness to random disruptions, it has become potentially more vulnerable to targeted shocks affecting central nodes—a finding with significant implications for supply chain risk management (Carvalho 2014; Bonadio et al. 2021).

6.3 | Correlation Patterns: Connecting Structure to Elasticities

The correlation analysis (Figures 4–6) provides a systematic mapping of relationships between network indicators and gravity coefficients across multiple aggregation levels.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

At the aggregate level (Figure 4), we observe strong negative correlations between distance coefficients and several network measures: assortativity ($r = -0.8$), global efficiency (GLE, $r = -0.8$), and various centralization indices. The near-perfect correlations ($|r| > 0.9$) between GDP coefficients and centralization measures (DZI, DZO, BZ) suggest that network concentration and economic mass effects are deeply intertwined—larger economies are both more central in the network and exhibit stronger GDP elasticities. Cells marked with a red '×' in Figure 4 indicate correlations that are not statistically significant at the 5% level after adjusting for multiple comparisons; the strong correlations discussed above all remain significant after this adjustment.

The sectoral correlation patterns (Figure A5) reveal substantial heterogeneity that would be masked by aggregate analysis. Manufacturing sectors (C10T12, C20, C24, C28) show distinct correlation profiles compared to service sectors (J58T60, J61, K), with manufacturing exhibiting stronger correlations between distance coefficients and out-degree centrality measures while services show stronger correlations with efficiency metrics. This heterogeneity extends to the cluster level (Figure 5), where NDA_1 sectors display moderate positive correlations between distance coefficients and lacunarity measures ($r \approx 0.4$), while NDA_2 and NDA_3 sectors show negative correlations.

Product-group correlations (Figure A6, Figure 6) exhibit even greater diversity. The correlation between distance elasticity and network resilience varies from strongly negative in primary products to near-zero in machinery and equipment. Figure 6a shows that clustered product groups display systematically different relationships with modularity and fractal dimension: NDA_1 products show weak positive correlations with MOD while NDA_2 products show negative correlations, suggesting that regional trade bloc formation affects product categories differentially.

These correlation patterns extend previous findings that network position affects trade outcomes (Fagiolo et al. 2009; Benedictis and Tajoli 2011) by demonstrating that the strength and even direction of network-gravity relationships varies systematically across sectors and products. This heterogeneity likely reflects differences in production fragmentation, transport cost structures, and the nature of demand (intermediate vs. final goods) across trade categories (Johnson and Noguera 2012; Koopman et al. 2014).

6.4 | Cross-Database Correspondence: Sectors and Products

The parallel analysis of OECD-ICIO sectors and BACI-CEPII product groups reveals meaningful correspondences that strengthen the robustness of our findings. Technology-intensive sectors in OECD-ICIO (C26 Computer and electronics, J61 Telecommunications) correspond to product groups 84–85 (Machinery, Electrical equipment) in BACI-CEPII; both exhibit similar patterns of declining distance sensitivity and strong correlations with network efficiency measures.

Specifically, we observe the following correspondences:

- *Services sectors* (J58T60, J61, K) in OECD-ICIO show NDA_1 clustering with declining $|\gamma|$; the corresponding service-related product groups in BACI (where trade in services is captured through embodied goods) exhibit analogous patterns.
- *Heavy manufacturing* (C24 Basic metals, C29 Motor vehicles) corresponds to product groups 72–76 (Base metals) and 87 (Vehicles), both showing NDA_2 or NDA_3 clustering with stable distance elasticity.
- *Primary sectors* (A01_02 Agriculture, B05_06 Mining) align with product groups 01–27 (Live animals, Mineral products), both exhibiting higher distance sensitivity and stronger correlations with resilience measures (RRes, SRes).

This cross-database consistency suggests that the observed network-gravity relationships are not artefacts of data construction but reflect genuine economic mechanisms operating across different trade conceptualizations (value-added vs. gross flows).

6.5 | Causal Relationships: From Correlation to Mechanism

The Granger causality analysis (Figures 7–10) addresses the causal component of our research question and constitutes a key contribution (C_2) by establishing directional temporal relationships using three complementary methods.

6.5.1 | Aggregate Network Causality

At the aggregate level, the traditional Granger causality results (Figure 7a) identify only four of the 29 network indicators as temporally preceding changes in distance elasticity: Power Centralization (PZ, lag 4), Average Path Length (AVPL, lag 5), Average Local Efficiency (ALE, lag 3), and Trace (lag 5). Neither modularity nor any of the four lacunarity moments nor fractal dimension shows a statistically significant relationship with the aggregate distance elasticity in this test. The causal network visualization (Figure 7b) shows the interdependencies among these few explanatory variables themselves.

For GDP elasticity (Figure 7c,d), the pattern is markedly different: 23 of the 29 network indicators, including all four lacunarity moments, fractal dimension, and the centralization measures PRZ and DZO, show a statistically significant traditional Granger relationship, almost uniformly at a one-year lag. This asymmetry—a small, longer-lagged set of predictors for distance elasticity versus a large, short-lagged set for GDP elasticity—is itself a substantive finding: it suggests that aggregate network reorganization translates into a broad, rapid adjustment of the mass (GDP) terms in the gravity equation, while its effect on the distance term is either genuinely weak at the aggregate level or is masked by offsetting, sector-specific dynamics that we examine directly in Section 6.5.3.

These findings suggest that structural changes in the trade network—particularly changes in centralization and in the heterogeneity of local structure (captured by lacunarity)—precede shifts in the GDP elasticities of the gravity equation more consistently than they precede shifts in the distance elasticity. This temporal ordering is consistent with theoretical models where

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

network reorganization (e.g., through preferential attachment) manifests rapidly in the mass terms of trade costs (Chaney 2014), while any corresponding effect on distance sensitivity appears to require the finer sectoral or product-level resolution examined below.

6.5.2 | Random Forest Results: Uncovering Nonlinear Relationships

The Random Forest-based causality analysis (Figure 8) yields importance rankings that differ notably from the traditional linear tests, revealing the presence of nonlinear causal mechanisms.

For distance elasticity (Figure 8a), the raw Random-Forest importance ranking differs from the traditional linear test, but none of the 29 indicators—including those with the highest raw importance—clears the shadow-feature significance threshold (Section Random Forest-Based Granger-Like Causality Analysis); we therefore do not interpret any indicator as a robust nonlinear predictor of the aggregate distance elasticity. For GDP elasticity (Figure 8b,c), by contrast, 23 of 29 indicators remain significant after the same shadow-feature correction for both origin and destination GDP, with Trace (domestic transactions/self-loops) the single strongest predictor in both directions by total variable importance—a nonlinear relationship that the traditional linear Granger test also detects (Section 5.4), but that the Random Forest test additionally confirms is not an artefact of a single functional-form assumption.

The systematic comparison of traditional, Bayesian, and Random Forest causality tests, once all three are evaluated against comparably strict significance criteria, reveals a consistent substantive pattern rather than a methodological artefact: modularity (MOD), harmonic closeness centralization (HZ), the Rich-Club Coefficient (RCC), and the fourth lacunarity moment (LL_m4) fail to robustly predict either GDP elasticity in any of the three methods, whereas Trace, PageRank centralization (PRZ), out-degree centralization (DZO), and the first two lacunarity moments (LL_m1, LL_m2) predict both GDP elasticities robustly across methods. This convergence across three methodologically distinct tests strengthens confidence in the GDP-elasticity findings specifically; it does not, by itself, strengthen the case for an aggregate network-to-distance-elasticity channel, which our data supports only at the sectoral and product level (Section 6.5.3).

The systematic comparison of traditional, Bayesian, and Random Forest causality tests reveals important patterns of convergence and divergence that inform interpretation:

- *Universal predictors*: Lacunarity moments (LL_m2, LL_m3) and path-related measures (SP, AVPL) are identified as significant predictors by all three methods, providing robust evidence for their causal role in gravity coefficient dynamics.
- *Method-specific findings*: Random Forest uniquely identifies Trace and Motifs as important predictors, suggesting nonlinear mechanisms involving domestic transaction patterns and local structural configurations that linear models cannot capture.
- *Bayesian confirmation*: The Bayesian approach, being more conservative, confirms approximately 70% of traditional Granger test results while providing continuous evidence measures that enable nuanced interpretation of marginal cases.

This methodological triangulation strengthens confidence in the core findings while revealing additional layers of complexity that merit further investigation.

6.5.3 | Panel Causality: Sectoral and Product Heterogeneity

The panel Granger causality analysis (Figures 9 and 10) extends the aggregate findings to reveal substantial heterogeneity across sectors and products.

For sectoral data (Figure 9), the traditional panel approach identifies Avg.PRC (average PageRank centrality), Avg.AC (authority centrality), and Avg.KC (Katz centrality) as having pooled causal effects on distance elasticity, with individual sector results showing considerable variation. Telecommunications (J61), paper products (C17_18), and arts/recreation (R) show the strongest individual causal patterns, while heavy industry sectors (C24–C29) show weaker or absent causality.

The Bayesian panel results (Figure 9b,e,h) provide a continuous measure of evidence strength that largely confirms the traditional results while offering a more nuanced assessment of marginal cases. The RF panel results (Figure 9c,f,i) again identify additional causal relationships, particularly for sectors where nonlinear dynamics may predominate.

For product groups (Figure 10), the pattern of causal heterogeneity is even more pronounced. Products in categories 53, 45, 59 (textiles and related products) show strong causal links from network centrality measures to distance elasticity, while machinery products (84, 85) show weaker causality. The Bayesian and RF approaches consistently identify more causal relationships than the traditional method, with RF finding significant nonlinear causality for product categories where linear tests fail.

6.6 | Theoretical Implications: From Network Topology to Trade Theory

Our Granger causality results have three principal theoretical implications that advance international trade theory beyond the methodological contributions.

6.6.1 | Extensive Margin Dynamics: Temporal-Precedence Evidence From Macro

The finding that network topology changes temporally precede gravity coefficient shifts by 1–4 years (confirming H_1 and H_2) provides systematic dynamic, macro-level evidence of temporal precedence consistent with the extensive margin mechanism proposed by heterogeneous firm models (Helpman et al. 2008; Melitz and Redding 2015; Chaney 2014); as with our other

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

causal-sounding claims, we describe this as Granger-style predictive evidence rather than structural causal proof (Section 8). While micro-level studies have documented the importance of extensive versus intensive margins in static decompositions (Helpman et al. 2008), the *temporal ordering* at the network level has not been established. Our results show that when trade costs change (due to technology, policy, or global shocks), the network topology adjusts first (reflecting firm entry/exit—the extensive margin), and only subsequently do aggregate gravity coefficients shift (reflecting volume adjustments—the intensive margin). The 1–4 year lag is economically meaningful: it reflects the time needed for firms' export market entry decisions to translate into measurable changes in aggregate bilateral trade patterns.

6.6.2 | A Structural Account of Sectoral Heterogeneity in the Distance Puzzle

Our sector-level analysis offers a novel, structural resolution to the long-debated distance puzzle (Disdier and Head 2008; Brun et al. 2005; Yotov 2012). The aggregate decline in distance elasticity masks two countervailing forces:

7. A *genuine* reduction in distance sensitivity for digitally tradable services and ICT products (NDA_1 cluster), driven by technology-enabled trade cost reduction (Lendle et al. 2016). These sectors exhibit rapidly declining modularity—trade links become geographically dispersed as digital technologies reduce the cost of serving distant markets.
8. *Persistent* distance sensitivity for physically intensive goods (NDA_3 cluster), consistent with Hummels and Klenow (2005)'s finding that transport costs for bulky goods have not declined proportionally. These sectors maintain high modularity—regional trade concentration persists because transport costs remain the binding constraint.

The network topology provides the *mechanism*: the Granger-causal link from modularity to distance elasticity (H_2) shows that the distance puzzle is fundamentally a *composition effect* operating through network reorganization. As the global trade network's composition shifts toward services (declining modularity, declining distance sensitivity), the aggregate distance elasticity falls—but this masks the persistent role of distance in goods trade. This resolution bridges the contradictory findings in the literature: Disdier and Head (2008)'s observation that distance remains important in meta-analysis is consistent with our goods-sector results, while Lendle et al. (2016)'s finding of declining distance effects applies specifically to our services/ICT clusters.

6.6.3 | Network Fragility and the Geography of Trade Costs

A novel theoretical insight emerges from the confirmation of H3: the strong causal relationship between lacunarity moments and gravity coefficients. Lacunarity captures the 'gappiness' of the trade partner distribution across geographic scales. When lacunarity increases—trade becomes more spatially heterogeneous—our results show that distance sensitivity subsequently rises. This is consistent with Baqaee and Farhi (2019)'s nonlinear amplification mechanism: as trade networks become more uneven in their geographic distribution, disruptions to concentrated hubs propagate disproportionately, effectively increasing the gravity of distance in trade.

This finding connects the gravity framework to the production network literature in a way not previously documented and has direct implications for the ongoing debate about geopolitical fragmentation (Alfaro and Chor 2023; Freund et al. 2024). The documented 1–4 year lag between network restructuring and gravity coefficient adjustment implies that the effects of recent reshoring and friendshoring initiatives on measured trade frictions may not be fully visible in gravity estimates for several years. Policymakers relying on current gravity estimates may therefore underestimate the eventual impact of supply chain reorganization policies. Our framework provides a tool for *nowcasting* these effects using network topology as a leading indicator. This distance-side finding is not the whole picture, however: the Random Forest results (Figure 8b,c) show that the local lacunarity moments (LL_m1 through LL_m4) and fractal dimension (FDim) are also among the more important predictors of both origin- and destination-GDP elasticities, consistent with the market-access channel outlined in Section 2. Because neither the Granger tests nor Random Forest importance scores identify the sign of this relationship, we treat whether structural heterogeneity ultimately strengthens or dampens the GDP elasticities as an open question for future research rather than infer a direction from significance alone.

6.7 | Relating Findings to Previous Research

Our results both confirm and extend previous findings in the gravity and network trade literature.

The declining aggregate distance elasticity confirms the pattern documented by Brun et al. (2005) and Disdier and Head (2008), while our disaggregated analysis extends their work by showing that this aggregate trend conceals substantial sectoral heterogeneity—some sectors show increasing distance sensitivity even as others decline. This reconciles apparently contradictory findings in the literature (Berthelon and Freund 2008; Yotov 2012) by showing that different sample compositions can yield different aggregate patterns.

Our finding that network centralization strongly correlates with GDP elasticity extends Fagiolo et al. (2009)'s observation that network centrality affects trade outcomes, demonstrating that this relationship operates not just cross-sectionally but through time and exhibits causal structure. The temporal precedence of network changes over gravity coefficient changes is a novel finding that suggests network analysis can provide leading indicators for shifts in fundamental trade parameters.

The identification of fractal dimension and lacunarity as predictors of gravity coefficients represents a genuinely new contribution. While García-Perez et al. (2016) connected fractal properties to trade network structure, we demonstrate that these measures have predictive content for the evolution of gravity parameters—a finding with both theoretical significance (suggesting scale-free properties matter for trade costs) and practical applications (providing early warning indicators for structural change).

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The systematic heterogeneity in network-gravity relationships across sectors extends the product-level gravity literature (Baldwin and Harrigan 2011; Hummels and Klenow 2005) by showing that not only do gravity coefficients differ across products, but so do the relationships between these coefficients and network structure. This finding suggests that one-size-fits-all approaches to modelling trade networks may miss important category-specific dynamics.

Our findings also speak to several recent developments in the trade literature. Antràs (2025) identifies the interaction between global sourcing decisions and network effects as one of the 'uncharted waters' of international trade theory. Our documentation of Granger-causal relationships between network topology and gravity parameters provides empirical evidence for exactly this interaction, showing that the network structure of global sourcing (captured by our multiplex trade networks) has measurable, temporally leading effects on the fundamental trade parameters. Similarly, Freund et al. (2024) document that US trade policy is already reshaping global supply chains; our framework provides a tool for monitoring how such policy-driven network restructuring will eventually manifest in measured trade frictions.

The growing literature on geopolitical fragmentation and 'friendshoring' (Alfaro and Chor 2023) raises the question of whether the decades-long decline in aggregate distance sensitivity will reverse. Our modularity-based analysis (H2) suggests that if trade network modularity increases—as friendshoring policies would imply—aggregate distance sensitivity should subsequently rise. The forecast trajectories in our analysis, which project continued but decelerating decline in $|\gamma_t|$ through 2030, may therefore underestimate future distance sensitivity if current geopolitical trends intensify.

7 | Conclusion

This study provides systematic evidence of temporal precedence—based on converging traditional, Bayesian, and machine-learning Granger-style tests, though not structural causal identification—linking the evolution of trade network topology to subsequent shifts in structural gravity coefficients, using two complementary databases spanning 1995–2022.

7.1 | Theoretical Contributions

Our findings advance international trade theory in three dimensions. *First, bridging gravity and network theories*. We establish a directional, temporal link between the structural gravity framework (Anderson and van Wincoop 2003; Yotov 2024) and production network theory (Acemoglu et al. 2012; Chaney 2014). The finding that network topology changes Granger-precede gravity coefficient shifts—but rarely vice versa—provides systematic dynamic evidence consistent with the extensive margin (network formation) leading the intensive margin (trade volumes) at the macroeconomic level, extending the static predictions of Melitz and Redding (2015) and Helpman et al. (2008) to a dynamic setting; as discussed in Section 8, this should be read as temporal-precedence evidence, not structural causal identification.

Second, a structural account of sectoral heterogeneity in the distance puzzle. Our most robust finding—consistent across the aggregate, sectoral, and product-group levels, and confirmed qualitatively by all three causality-testing methods—is that the distance puzzle (Disdier and Head 2008; Brun et al. 2005; Yotov 2012) is not a singular phenomenon but a composition effect: ICT and services sectors exhibit genuine distance compression through network decentralization, while heavy manufacturing maintains geographic constraints through persistent regional modularity. This sector-level decomposition, grounded in the Granger-style temporal-precedence link between modularity and distance elasticity, offers a structural, testable account of why aggregate estimates of the distance puzzle have proven so heterogeneous across studies and samples, though we stop short of claiming it fully resolves the underlying debate.

Third, linking network fragility to trade frictions. Our finding that lacunarity moments and fractal dimension Granger-precede gravity coefficient shifts connects the growing literature on trade network resilience (Bonadio et al. 2021; Carvalho et al. 2021) to the gravity framework. At the aggregate network level, this heterogeneity-to-gravity link is clearest for GDP elasticities: three of the four lacunarity moments and fractal dimension Granger-precede both origin and destination GDP elasticity, consistent with nonlinear shock amplification mechanisms (Baqaee and Farhi 2019). For distance elasticity, the aggregate-level link is weaker—none of the four lacunarity moments or fractal dimension Granger-precede the aggregate distance coefficient—and, consistent with our second contribution, instead emerges at the sectoral and product-group level. This provides both a theoretical mechanism and an empirical tool for understanding how structural vulnerabilities translate into measurable changes in trade frictions, operating principally through economic-mass elasticities at the aggregate level and through distance elasticities at a more granular, sector-specific level.

7.2 | Methodological Contributions

Methodologically, we contribute a triangulated causality framework that combines traditional panel Granger tests, Bayesian model comparison, and Random Forest-based variable importance analysis. The machine learning approach reveals nonlinear causal relationships (particularly the role of Trace and Motifs) that linear methods entirely miss, underscoring the value of methodological pluralism in network-gravity analysis.

7.3 | Policy Implications

Our findings carry several actionable implications for trade policy design and implementation.

7.3.1 | Early Warning System Design

The documented 1–4 year lead times between network indicator changes and gravity coefficient shifts enable construction of operational early warning systems. We recommend monitoring:

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

- **Monthly/Quarterly**: PageRank centralization (PRZ) and modularity (MOD) as short-term (1–2 year) leading indicators.
- **Annually**: Lacunarity third moment (LL_m3) and fractal dimension (FDim) as medium-term (2–4 year) structural change indicators.
- **Threshold alerts**: SRes falling below RRes as a vulnerability signal.

7.3.2 | Sector-Specific Policy Targeting

The substantial heterogeneity documented across sectors implies differentiated policy approaches:

- **Digital/ICT sectors** (J61, J62_63): Prioritize connectivity and efficiency improvements; distance barriers are already low and declining further.
- **Heavy manufacturing** (C24–C29): Focus on partner diversification rather than distance reduction; resilience measures show these sectors are vulnerable to hub disruption.
- **Agricultural and primary sectors**: Monitor resilience indicators closely; these sectors show the strongest coupling between network fragility and distance sensitivity.

7.3.3 | Regional Trade Agreement Evaluation

The rising modularity trend (Figure 3a) combined with its Granger-causal relationship to distance elasticity suggests that trade agreements are not merely redirecting flows but fundamentally restructuring trade costs. Policy evaluation frameworks should incorporate network topology changes as endogenous outcomes rather than treating gravity coefficients as stable parameters.

7.3.4 | Supply Chain Resilience Assessment

The documented divergence between random (RRes) and systematic (SRes) attack resilience provides a quantitative basis for identifying vulnerable sectors. Sectors where $\text{SRes} < 0.95 \times \text{RRes}$ should be flagged for supply chain diversification initiatives, as they face elevated risk from targeted disruptions (sanctions, geopolitical tensions) even if robust to random shocks.

8 | Limitations and Future Work

Several limitations should be acknowledged. First, a single common shock would be expected to move all sectors similarly, whereas we document systematic sector- and product-level heterogeneity in the estimated lead–lag pattern (Section 6.5.3); this heterogeneity is harder, though not impossible, to reconcile with a single common-shock story than with a sector-specific network-to-gravity channel. We report (Section Rubustness Tests of Random-Forest Causality Methods and Placebo Tests for Granger Causality) a false-discovery-rate-corrected and placebo-benchmarked version of our panel Granger results. After Benjamini–Hochberg correction (Benjamini and Hochberg 1995) at the individual sector level (1008 sector-by-indicator tests per coefficient), only 0 of 39 raw-significant distance relationships and 0 of 44 raw-significant destination-GDP relationships survive at $q = 0.05$ (2 of 45 for origin GDP); at the pooled panel level (12 sector-fixed-effects tests per coefficient), by contrast, 11 of 12 origin-GDP and 8 of 12 destination-GDP tests remain significant at $q = 0.05$. A complementary circular-shift placebo test (Theiler et al. 1992) shows that the observed number of significant sector-level pairs (117 for distance, 137 for origin GDP, 130 for destination GDP) exceeds the 95th percentile of its placebo null distribution in every case (99, 86, and 114, respectively; empirical $p < 0.01$ for all three) (Table A4). Taken together, these diagnostics indicate that individual sector-by-indicator claims should be treated cautiously—most do not survive correction for testing many sector/variable/lag combinations—but that the pooled, aggregate-level pattern of network-to-gravity predictability is not solely an artefact of multiple testing. By construction, neither diagnostic can distinguish a genuine network-to-gravity mechanism from an unobserved common driver with a similar lag structure; we therefore describe our results throughout as evidence of robust temporal precedence rather than as establishing a causal mechanism (see the revised framing in Section 1 and the Conclusion). Future work could employ instrumental-variable approaches (e.g., plausibly exogenous shocks to specific trade corridors) or quasi-experimental designs (e.g., staggered regional-trade-agreement entry into force) to strengthen causal identification.

Second, the relatively short time series (26–28 annual observations) limits the power of causality tests, increases sensitivity to lag-order choice, and raises the risk of spurious findings when many sector/product/lag combinations are tested; it also limits the ability to detect long-horizon relationships or multiple structural breaks. We now report, alongside our main results, a false-discovery-rate-corrected version of the panel Granger tests and a circular-shift placebo benchmark that directly quantify this risk rather than relying on cross-method triangulation alone. Triangulation across traditional, Bayesian, and machine-learning approaches remains informative about the robustness of a finding to modelling choices, but, as the reviewer correctly notes, it cannot by itself overcome the fundamental power limitation of a 26–28-year panel. As longer series become available, extended analysis could identify regime-specific patterns and long-run equilibrium relationships with adequately powered tests.

Third, our analysis treats sectors and product groups as independent units, ignoring input–output linkages that may propagate network effects across categories. Future work integrating the multiplex structure of inter-sectoral trade could reveal how network changes in upstream sectors cascade to downstream gravity coefficients.

Fourth, we focus on aggregate country-level networks, abstracting from firm-level heterogeneity that may be important for understanding micro-level mechanisms. Firm-level trade data, where available, could enable analysis of how the network position of individual firms relates to their distance sensitivity.

Fifth, structural gravity specification. As explained in Section 4.1.3, a specification with country-pair fixed effects is not compatible with our object of interest—year-specific distance and GDP

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

elasticities—because distance is time-invariant per pair and a pair fixed effect would saturate a design with one observation per pair per year; exporter- or importer-year fixed effects are similarly infeasible here, because within a single year they would be collinear with the very GDP elasticities $\beta_{o,t}$ and $\beta_{d,t}$ that we aim to estimate. Instead, Equation (1) includes nine bilateral controls from the CEPII Gravity database (Conte et al. 2022) directly in every year-specific cross-section: common official and ethnic language, contiguity, colonial history, common legal origin, joint FTA/WTO membership, and EU membership of either partner (Section 4.1.3). These bilateral controls absorb considerable time-invariant bilateral heterogeneity, but, as discussed in Section 3.2, they cannot substitute for the country-specific multilateral resistance terms that exporter- or importer-year fixed effects would otherwise capture. This specification also still falls short of the full structural-gravity toolkit recommended by Larch et al. (2025) for panel gravity estimation, and it does not explicitly model firm-level heterogeneity (Helpman et al. 2008; Melitz and Redding 2015). Future work should extend our framework to incorporate the full structural gravity toolkit, potentially using the constrained PPML approach of Larch et al. (2025) to separate genuine distance-elasticity evolution from changes in correlated bilateral factors.

Sixth, macro-level versus firm-level analysis. Our analysis operates at the country-sector and country-product level, which necessarily aggregates over firm-level heterogeneity. While Chaney (2014) provides the theoretical link between firm-level decisions and network topology, our data do not allow direct observation of the firm-level mechanisms underlying the documented Granger-causal relationships. Future work using firm-level customs data could test whether the extensive margin dynamics we identify at the macro level are indeed driven by firm entry/exit patterns.

Finally, future work should explore the forecasting value of network indicators more formally through out-of-sample prediction exercises, and investigate whether the identified relationships remain stable during periods of major disruption (such as the COVID-19 pandemic) or exhibit structural breaks that would limit their predictive utility.

Funding

This work was supported by OTKA, K 142395.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available at https://github.com/kzst/GRAV_ICIO_CEPII.

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi. 2012. "The Network Origins of Aggregate Fluctuations." *Econometrica* 80: 1977–2016. https://doi.org/10.3982/ECTA9623.

Alfaro, L., and D. Chor. 2023. "Global Supply Chains: The Looming 'Great Reallocation'." Working Paper 31661. National Bureau of Economic Research. https://doi.org/10.3386/w31661.

Anderson, J. E. 1979. "A Theoretical Foundation for the Gravity Equation." *American Economic Review* 69: 106–116. https://doi.org/10.2307/1802501.

Anderson, J. E. 2011. "The Gravity Model." *Annual Review of Economics* 3: 133–160. https://doi.org/10.1146/annurev-economics-111809-125114.

Anderson, J. E., and E. van Wincoop. 2003. "Gravity With Gravitas: A Solution to the Border Puzzle." *American Economic Review* 93: 170–192. https://doi.org/10.1257/000282803321455214.

Anderson, J. E., and Y. V. Yotov. 2010. "The Changing Incidence of Geography." *American Economic Review* 100: 2157–2186. https://doi.org/10.1257/aer.100.5.2157.

Antràs, P. 2025. "The Uncharted Waters of International Trade." *Journal of the European Economic Association* 23: 1–58. https://doi.org/10.1093/jeea/jvae060.

Antràs, P., and D. Chor. 2013. "Organizing the Global Value Chain." *Econometrica* 81: 2127–2204. https://doi.org/10.3982/ECTA9828.

Arkolakis, C., A. Costinot, and A. Rodríguez-Clare. 2012. "New Trade Models, Same Old Gains?" *American Economic Review* 102: 94–130. https://doi.org/10.1257/aer.102.1.94.

Baier, S. L., and J. H. Bergstrand. 2007. "Do Free Trade Agreements Actually Increase Members' International Trade?" *Journal of International Economics* 71: 72–95. https://doi.org/10.1016/j.jinteco.2006.02.005.

Baldwin, R., and J. Harrigan. 2011. "Zeros, Quality, and Space: Trade Theory and Trade Evidence." *American Economic Review* 101: 2920–2940. https://doi.org/10.1257/aer.101.7.2920.

Baldwin, R., and D. Taglioni. 2006. "Gravity for Dummies and Dummies for Gravity in International Economics." NBER Working Paper 12516. National Bureau of Economic Research. https://doi.org/10.3386/w12516.

Baqaee, D. R., and E. Farhi. 2019. "The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem." *Econometrica* 87: 1155–1203. https://doi.org/10.3982/ECTA15202.

Barigozzi, M., G. Fagiolo, and D. Garlaschelli. 2010. "Multinetwork of International Trade: A Commodity-Specific Analysis." *PLoS One* 5: e12249. https://doi.org/10.1371/journal.pone.0012249.

Benedictis, L. D., and L. Tajoli. 2011. "The World Trade Network." *World Economy* 34: 1417–1454. https://doi.org/10.1111/j.1467-9701.2011.01360.x.

Benjamini, Y., and Y. Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society Series B, Statistical Methodology* 57: 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Bergmeir, C., R. J. Hyndman, and B. Koo. 2018. "A Note on the Validity of Cross-Validation for Evaluating Autoregressive Time Series Prediction." *Machine Learning* 108: 349–355. https://doi.org/10.1007/s10994-018-5711-1.

Bernard, A. B., and A. Moxnes. 2018. "Networks and Trade." *Annual Review of Economics* 10: 65–85. https://doi.org/10.1146/annurev-economics-080217-053506.

Berthelon, M., and C. Freund. 2008. "On the Conservation of Distance in International Trade." *Journal of International Economics* 75: 310–320. https://doi.org/10.1016/j.jinteco.2008.06.002.

Boehm, C. E., A. Flaaen, and N. Pandalai-Nayar. 2019. "Input Linkages and the Transmission of Shocks: Firm-Level Evidence From the 2011 tōhoku Earthquake." *Quarterly Journal of Economics* 134: 125–173. https://doi.org/10.1093/qje/qjz025.

Bonadio, B., Z. Huo, A. A. Levchenko, and N. Pandalai-Nayar. 2021. "Global Supply Chains in the Pandemic." *Journal of International Economics* 133: 103391. https://doi.org/10.1016/j.jinteco.2021.103391.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45: 5–32. https://doi.org/10.1023/A:1010933404324.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Brun, J. F., C. Carrère, P. Guillaumont, and J. de Melo. 2005. "Has Distance Died? Evidence From a Panel Gravity Model." *World Bank Economic Review* 19: 99–120. https://doi.org/10.1093/wber/lhi004.

Caliendo, L., and F. Parro. 2015. "Estimates of the Trade and Welfare Effects of NAFTA." *Review of Economic Studies* 82: 1–44. https://doi.org/10.1093/restud/rdu035.

Carrère, C., M. Mrázová, and J. P. Neary. 2020. "Gravity Without Apology: The Science of Elasticities, Distance, and Trade." *Economic Journal* 130: 880–910. https://doi.org/10.1093/ej/ueaa034.

Carvalho, V. M. 2014. "From Micro to Macro via Production Networks." *Journal of Economic Perspectives* 28: 23–48. https://doi.org/10.1257/jep.28.4.23.

Carvalho, V. M., M. Nirei, Y. U. Saito, and A. Tahbaz-Salehi. 2021. "Supply Chain Disruptions: Evidence From the Great East Japan Earthquake." *Quarterly Journal of Economics* 136: 1255–1321. https://doi.org/10.1093/qje/qjab012.

Cerina, F., Z. Zhu, A. Chessa, and M. Riccaboni. 2015. "World Input-Output Network." *PLoS One* 10: e0134025. https://doi.org/10.1371/journal.pone.0134025.

Chaney, T. 2014. "The Network Structure of International Trade." *American Economic Review* 104: 3600–3634. https://doi.org/10.1257/aer.104.11.3600.

Chen, B., H. Xu, and L. Feng. 2023. "Structural Properties and Evolution of Global Photovoltaic Industry Trade Network." *Environmental Science and Pollution Research* 30: 69580–69598. https://doi.org/10.1007/s11356-023-27213-9.

Chudik, A., and M. H. Pesaran. 2015. "Common Correlated Effects Estimation of Heterogeneous Dynamic Panel Data Models With Weakly Exogenous Regressors." *Journal of Econometrics* 188: 393–420. https://doi.org/10.1016/j.jeconom.2014.10.011.

Chung, M. G., K. Kapsar, K. A. Frank, and J. Liu. 2020. "The Spatial and Temporal Dynamics of Global Meat Trade Networks." *Scientific Reports* 10: 16657. https://doi.org/10.1038/s41598-020-73591-2.

Conte, M., P. Cotterlaz, and T. Mayer. 2022. "The CEPII Gravity Database." CEPII Working Paper 2022-05. CEPII. Paris. https://www.cepii.fr/CEPII/en/publications/wp/abstract.asp?NoDoc=13432.

di Giovanni, J., A. A. Levchenko, and I. Méjean. 2020. "The Micro Origins of International Business Cycle Comovement." *Review of Economic Studies* 87: 1339–1383. https://doi.org/10.1093/restud/rdaa021.

Ding, S., L. Wang, and Q. Zhou. 2025. "Study on Evolution Mechanism of Agricultural Trade Network of Rcep Countries—Complex System Analysis Based on the Tergm Model." *Systems* 13: 593. https://doi.org/10.3390/systems13070593.

Disdier, A. C., and K. Head. 2008. "The Puzzling Persistence of the Distance Effect on Bilateral Trade." *World Economy* 31: 96–118. https://doi.org/10.1111/j.1467-9701.2007.01010.x.

Dong, Q., J. Zhou, and Q. Du. 2024. "Analysis of the Spatial Correlation Pattern of Logistics Carbon Emission Efficiency and Its Influencing Factors: The Case of China." *Environmental Science and Pollution Research* 31: 11178–11191. https://doi.org/10.1007/s11356-023-31753-5.

Dueñas, M., and A. Mandel. 2023. "The Structure of Global Cultural Networks: Evidence From the Diffusion of Music Videos." *PLoS One* 18: e0294149. https://doi.org/10.1371/journal.pone.0294149.

Dumitrescu, E. I., and C. Hurlin. 2012. "Testing for Granger Non-Causality in Heterogeneous Panels." *Economic Modelling* 29: 1450–1460. https://doi.org/10.1016/j.econmod.2012.02.014.

Eaton, J., and S. Kortum. 2002. "Technology, Geography, and Trade." *American Economic Review* 92: 170–192. https://doi.org/10.1257/000282802762024502.

Egger, P., and S. Nigai. 2015. "Structural Gravity With Multilateral Resistance Terms at Work." *Journal of International Economics* 96: 206–217. https://doi.org/10.1016/j.jinteco.2015.02.008.

Egger, P., and M. Pfaffermayr. 2003. "The Proper Panel Econometric Specification of the Gravity Equation: A Three-Way Model With Bilateral Interaction Effects." *Empirical Economics* 28: 571–580. https://doi.org/10.1007/s00181-003-0182-0.

Fagiolo, G., J. Reyes, and S. Schiavo. 2009. "World-Trade Web: Topological Properties, Dynamics, and Evolution." *Physica A: Statistical Mechanics and Its Applications* 387: 3868–3873. https://doi.org/10.1016/j.physa.2008.12.020.

Fally, T. 2015. "Structural Gravity and Fixed Effects." *Journal of International Economics* 97: 76–85. https://doi.org/10.1016/j.jinteco.2015.05.005.

Fang, Y., M. Tang, H. Wei, Z. Feng, and N. Yu. 2023. "Global Ginseng Trade Networks: Structural Characteristics and Influencing Factors." *Frontiers in Pharmacology* 14: 1119183. https://doi.org/10.3389/fphar.2023.1119183.

Feng, L., B. Chen, G. Wu, and Q. Zhang. 2024. "Global Renewable Energy Trade Network: Patterns and Determinants." *Environmental Science and Pollution Research* 31: 15538–15558. https://doi.org/10.1007/s11356-024-32066-x.

Flori, A., and A. Spelta. 2025. "Carbon Trade Biases and the Emerging Mesoscale Structure of the European Emissions Trading System Network." *Nature Communications* 16: 5199. https://doi.org/10.1038/s41467-025-59913-w.

Freund, C., A. Mattoo, A. Mulabdic, and M. Ruta. 2024. "Is US Trade Policy Reshaping Global Supply Chains?" *Journal of International Economics* 152: 104011. https://doi.org/10.1016/j.jinteco.2024.104011.

García-Perez, G., M. Boguñá, A. Allard, and M. Serrano. 2016. "The Hidden Hyperbolic Geometry of International Trade: World Trade Atlas 1870–2013." *Scientific Reports* 6: 1–10. https://doi.org/10.1038/srep33441.

Garlaschelli, D., and M. I. Loffredo. 2004. "Fitness-Dependent Topological Properties of the World Trade Web." *Physical Review Letters* 93: 188701. https://doi.org/10.1103/PhysRevLett.93.188701.

Gaulier, G., and S. Zignago. 2010. "BACI: International Trade Database at the Product-Level." CEPII Working Paper 2010-23. CEPII. https://doi.org/10.2139/ssrn.1994502.

Granger, C. W. J. 1969. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods." *Econometrica* 37: 424–438. https://doi.org/10.2307/1912791.

Hao, S., A. Kassahun, Y. Bouzembrak, and H. Marvin. 2020. "Identification of Potential Vulnerable Points and Paths of Contamination in the Dutch Broiler Meat Trade Network." *PLoS One* 15: e0233376. https://doi.org/10.1371/journal.pone.0233376.

He, S., X. Dai, and C. He. 2024. "Symphony of Sustainability: How Technology Similarity Impacts International Trade Patterns in Environmental Goods." *Environmental Science and Pollution Research* 31: 23747–23765. https://doi.org/10.1007/s11356-024-32634-1.

Head, K., and T. Mayer. 2014. "Gravity Equations: Workhorse, Toolkit, and Cookbook." In *Handbook of International Economics*, edited by G. Gopinath, E. Helpman, and K. Rogoff, vol. 4, 131–195. Elsevier. https://doi.org/10.1016/B978-0-444-54314-1.00003-3.

Helpman, E., M. Melitz, and Y. Rubinstein. 2008. "Estimating Trade Flows: Trading Partners and Trading Volumes." *Quarterly Journal of Economics* 123: 441–487. https://doi.org/10.1162/qjec.2008.123.2.441.

Hidalgo, C. A., and R. Hausmann. 2009. "The Building Blocks of Economic Complexity." *Proceedings of the National Academy of Sciences* 106: 10570–10575. https://doi.org/10.1073/pnas.0900943106.

Holtz-Eakin, D., W. Newey, and H. S. Rosen. 1988. "Estimating Vector Autoregressions With Panel Data." *Econometrica* 56: 1371–1395. https://doi.org/10.2307/1913103.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Hood, L., and J. Gassner. 2017. "Gravity and the Global Network of International Trade." *Journal of Complex Networks* 5: 656–679. https://doi.org/10.1093/comnet/cnx006.

Hummels, D., and P. J. Klenow. 2005. "The Variety and Quality of a Nation's Exports." *American Economic Review* 95: 704–723. https://doi.org/10.1257/0002828054201396.

Jeffreys, H. 1961. *Theory of Probability*. 3rd ed. Oxford University Press.

Jiang, Y., Y. Li, Y. Li, Y. Xu, and E. Veglianti. 2023. "Research on the Structural Characteristics and Influencing Factors of Global Environmental Services Trade Networks." *Environmental Science and Pollution Research* 30: 53063–53076. https://doi.org/10.1007/s11356-023-26152-9.

Johnson, R. C., and G. Noguera. 2012. "Accounting for Intermediates: Production Sharing and Trade in Value Added." *Journal of International Economics* 86: 224–236. https://doi.org/10.1016/j.jinteco.2012.05.002.

Kass, R. E., and A. E. Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90: 773–795. https://doi.org/10.1080/01621459.1995.10476572.

Kivelä, M., A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter. 2014. "Multilayer Networks." *Journal of Complex Networks* 2: 203–271. https://doi.org/10.1093/comnet/cnu016.

Koop, G., and D. Korobilis. 2010. "Bayesian Multivariate Time Series Methods for Empirical Macroeconomics." *Foundations and Trends in Econometrics* 3: 267–358. https://doi.org/10.1561/0800000013.

Koop, G., and D. Korobilis. 2013. "Large Time-Varying Parameter VARs." *Journal of Econometrics* 177: 185–198. https://doi.org/10.1016/j.jeconom.2013.04.007.

Koopman, R., Z. Wang, and S. J. Wei. 2014. "Tracing Value-Added and Double Counting in Gross Exports." *American Economic Review* 104: 459–494. https://doi.org/10.1257/aer.104.2.459.

Kosowska-Stamirowska, Z., C. Ducruet, and N. Rai. 2016. "Evolving Structure of the Maritime Trade Network: Evidence From the lloyd's Shipping Index (1890–2000)." *Journal of Shipping and Trade* 1: 10. https://doi.org/10.1186/s41072-016-0013-3.

Kosztyán, Z. T., A. I. Katona, M. T. Kurbucz, and Z. Lantos. 2024. "Generalized Network-Based Dimensionality Analysis." *Expert Systems With Applications* 238: 121779. https://doi.org/10.1016/j.eswa.2023.121779.

Kursa, M. B., and W. R. Rudnicki. 2010. "Feature Selection With the Boruta Package." *Journal of Statistical Software* 36: 1–13. https://doi.org/10.18637/jss.v036.i11.

Larch, M., S. Shikher, and Y. V. Yotov. 2025. "Estimating Gravity Equations: Theory Implications, Econometric Developments, and Practical Recommendations." *Review of International Economics* 33: 1066–1092. https://doi.org/10.1111/roie.12789.

Lee, D., W. Cho, H. Kim, G. Kim, H. C. Jeong, and B. J. Kim. 2025. "Self-Consistent Gravity Model for Inferring Node Mass in Flow Networks." *Scientific Reports* 15: 18839. https://doi.org/10.1038/s41598-025-03664-7.

Lendle, A., M. Olarreaga, S. Schropp, and P. L. Vézina. 2016. "There Goes Gravity: Ebay and the Death of Distance." *American Economic Review* 106: 2905–2919. https://doi.org/10.1257/aer.20140143.

Liang, F., R. Paulo, G. Molina, M. A. Clyde, and J. O. Berger. 2008. "Mixtures of g Priors for Bayesian Variable Selection." *Journal of the American Statistical Association* 103: 410–423. https://doi.org/10.1198/016214507000001337.

Mayer, T., and S. Zignago. 2011. "Notes on CEPII's Distances Measures: The GeoDist Database." CEPII Working Paper 2011-25. CEPII. https://doi.org/10.2139/ssrn.1994531.

Melitz, M. J., and S. J. Redding. 2015. "New Trade Models, New Welfare Implications." *American Economic Review* 105: 1105–1146. https://doi.org/10.1257/aer.20130351.

Moritz, S., and T. Bartz-Beielstein. 2017. "imputeTS: Time Series Missing Value Imputation in R." *R Journal* 9: 207–218. https://doi.org/10.32614/RJ-2017-009.

Novy, D. 2013. "Gravity Redux: Measuring International Trade Costs With Panel Data." *Economic Journal* 123: F535–F559. https://doi.org/10.1111/ecoj.12019.

Opsahl, T., F. Agneessens, and J. Skvoretz. 2010. "Node Centrality in Weighted Networks: Generalizing Degree and Shortest Paths." *Social Networks* 32: 245–251. https://doi.org/10.1016/j.socnet.2010.03.006.

Pesaran, M. H. 2006. "Estimation and Inference in Large Heterogeneous Panels With a Multifactor Error Structure." *Econometrica* 74: 967–1012. https://doi.org/10.1111/j.1468-0262.2006.00692.x.

Poon, J. P. H., P. Peng, and J. D. Atkinson. 2024. "Industrial and Textile Waste Trade: Multilayer Network and Environmental Policy Effects." *Waste Management* 177: 146–157. https://doi.org/10.1016/j.wasman.2024.01.048.

Robinson, P., A. McInnes, N. Sajedianfard, M. Melatos, and J. A. Henderson. 2025. "Electrical Circuit Model of Spatiotemporal Trade Dynamics: Foundations and Derivation of the Gravity Model." *PLoS One* 20: e0326102. https://doi.org/10.1371/journal.pone.0326102.

Rouder, J. N., R. D. Morey, P. L. Speckman, and J. M. Province. 2012. "Default Bayes Factors for ANOVA Designs." *Journal of Mathematical Psychology* 56: 356–374. https://doi.org/10.1016/j.jmp.2012.08.001.

Ryczkowski, T., A. Fronczak, and P. Fronczak. 2017. "How Transfer Flights Shape the Structure of the Airline Network." *Scientific Reports* 7: 5630. https://doi.org/10.1038/s41598-017-06108-z.

Santos Silva, J., and S. Tenreyro. 2006. "The Log of Gravity." *Review of Economics and Statistics* 88: 641–658. https://doi.org/10.1162/rest.88.4.641.

Serrano, M. Á., and M. Boguñá. 2003. "Topological Properties of the World Trade Web: A Weighted Network Analysis." *Physical Review E* 68: 015101(R). https://doi.org/10.1103/PhysRevE.68.015101.

Sgrignoli, P., R. Metulini, S. Schiavo, and M. Riccaboni. 2015. "The Relation Between Global Migration and Trade Networks." *Physica A: Statistical Mechanics and Its Applications* 417: 245–260. https://doi.org/10.1016/j.physa.2014.09.037.

Silva, J. M. C. S., and S. Tenreyro. 2011. "Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator." *Economics Letters* 112: 220–222. https://doi.org/10.1016/j.econlet.2011.05.008.

Simonovska, I., and M. E. Waugh. 2014. "The Elasticity of Trade: Estimates and Evidence." *American Economic Review* 104: 1123–1148. https://doi.org/10.1257/aer.104.4.1123.

Song, C., S. Havlin, and H. A. Makse. 2005. "Self-Similarity of Complex Networks." *Nature* 433: 392–395. https://doi.org/10.1038/nature03248.

Theiler, J., S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer. 1992. "Testing for Nonlinearity in Time Series: The Method of Surrogate Data." *Physica D, Nonlinear Phenomena* 58: 77–94. https://doi.org/10.1016/0167-2789(92)90102-S.

Timmer, M., A. A. Erumban, G. de Vries, R. Stehrer, and B. Los. 2015. "An Illustrated User Guide to the World Input–Output Database: The Case of Global Automotive Production." *Review of International Economics* 23: 575–605. https://doi.org/10.1111/roie.12178.

Vespignani, A. 2012. "Modelling Dynamical Processes in Complex Socio-Technical Systems." *Nature Physics* 8: 32–39. https://doi.org/10.1038/nphys2160.

Wang, Y., and S. Cui. 2025. "Deepening Digital Intellectual Property Rules and Digital International Patenting." *Economics Letters* 251: 112312. https://doi.org/10.1016/j.econlet.2025.112312.

Xing, L., and W. Chen. 2023. "Structural Characteristics and Evolutionary Drivers of Global Virtual Water Trade Networks: A

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Stochastic Actor-Oriented Model for 2000–2015." *International Journal of Environmental Research and Public Health* 20: 3234. https://doi.org/10.3390/ijerph20043234.

Yotov, Y. V. 2012. "A Simple Solution to the Distance Puzzle in International Trade." *Economics Letters* 117: 794–798. https://doi.org/10.1016/j.econlet.2012.05.032.

Yotov, Y. V. 2024. "The Evolution of Structural Gravity: The Workhorse Model of Trade." *Contemporary Economic Policy* 42: 578–603. https://doi.org/10.1111/coep.12666.

Appendix A: Technical Appendix

Forecasting of Elasticies

Figure A1 presents the GNDA-identified clusters of industrial sectors based on the temporal similarity of their gravity coefficients, with forecasts to 2030. The clustering uses Spearman correlation as the similarity measure.

For distance coefficients (Figure A1a), three clusters emerge: NDA_1 (23 sectors) includes service-oriented and domestic-focused industries such as agriculture (A01_02), accommodation (I), real estate (L), and public services (O, P, Q), exhibiting distance coefficients around − 0.55 to − 0.45; NDA_2 (12 sectors) comprises resource-intensive and manufacturing industries including mining (B05_06, B07_08), food products (C10T12), chemicals (C20), and wholesale trade (G), with coefficients between − 0.50 and − 0.45; NDA_3 (9 sectors) contains technology-intensive manufacturing such as paper products (C17_18), electronics (C26), electrical equipment (C27), and telecommunications (J61), showing the steepest decline in distance sensitivity from − 0.60 to − 0.45.

For origin GDP coefficients (Figure A1b), NDA_1 (26 sectors) groups export-capacity-sensitive industries including chemicals (C20), machinery (C28), and professional services (M); NDA_2 (10 sectors) contains labour-intensive sectors such as textiles (C13T15), utilities (D, E), and education (P); NDA_3 (8 sectors) includes capital-intensive industries like basic metals (C24), electronics (C26), and land transport (H49).

For destination GDP coefficients (Figure A1c), NDA_1 (24 sectors) encompasses market-access-sensitive industries including most manufacturing sectors and business services; NDA_2 (13 sectors) contains domestically-oriented services such as accommodation (I), real estate (L), and public administration (O); NDA_3 (7 sectors) groups transport and electronics sectors including water transport (H50), warehousing (H52), and motor vehicles (C29).

Figure A2 displays the GNDA-identified clusters of product groups (HS-2 level, 97 categories) from the CEPII-BACI database, based on temporal patterns in gravity coefficients.

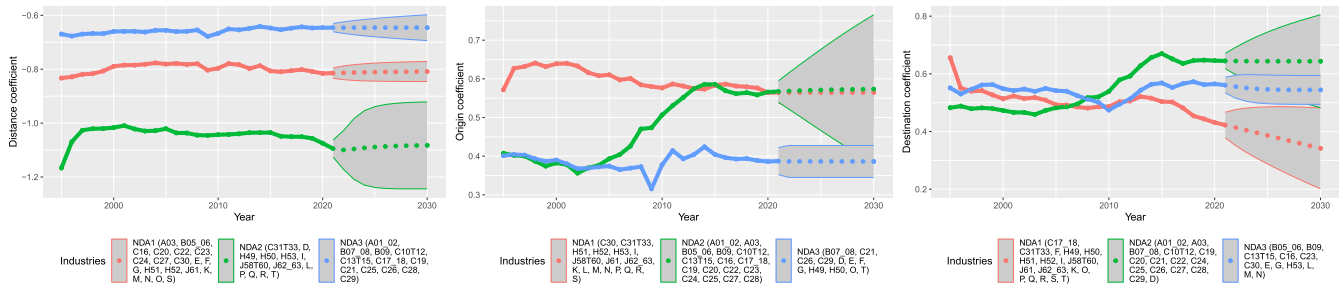
For distance coefficients (Figure A2a), three clusters are identified: NDA_1 (49 product groups) includes manufactured goods such as vehicles (87), machinery (84), electrical equipment (85), textiles (55–59), and footwear (64), showing distance coefficients stabilizing around − 0.10; NDA_2 (41 product groups) comprises primary products and intermediates including live animals (01), minerals (25–27), base metals (72–76), and paper (48), with coefficients around − 0.15; NDA_3 (6 product groups) contains specific categories including fertilizers (28, 29), cork products (53, 54), lead (78), and ships (89), exhibiting the highest distance sensitivity near − 0.20.

For origin GDP coefficients (Figure A2b), NDA_1 (47 groups) includes differentiated manufactures; NDA_2 (25 groups) contains agricultural and food products; NDA_3 (24 groups) comprises raw materials and standardized intermediates.

For destination GDP coefficients (Figure A2c), NDA_1 (61 groups) encompasses most consumer and industrial goods; NDA_2 (35 groups) includes commodities and specialized equipment.

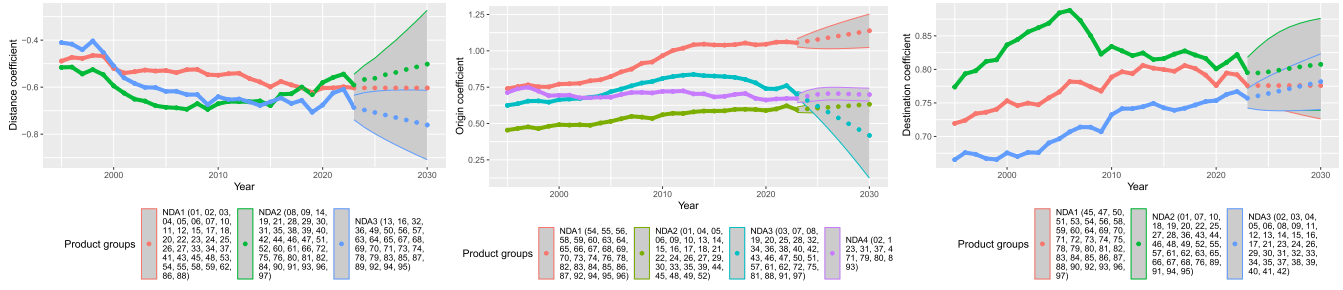
Figure A3 presents the clustering results for product subgroups (HS-4 level, approximately 1200 categories) from the CEPII-BACI database.

For distance coefficients (Figure A3a), three clusters emerge with coefficients ranging from − 0.30 (NDA_3) to − 0.10 (NDA_1). NDA_1 (49 subgroups) includes specialized products such as dairy (0401), nuclear reactors (8401–8402), and motorcycles (8711); NDA_2 (approximately 350 subgroups) contains diverse intermediates including live animals,



(a) Distance (b) Origin (c) Destination

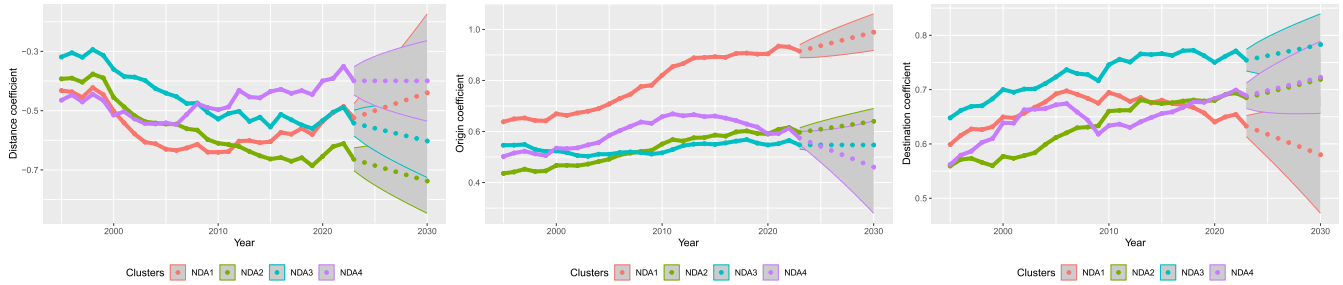
FIGURE A1 | Coefficients of gravity models in time (forecast wth ARIMA, 95% confidence interval, based on OECD-ICIO database). [Colour figure can be viewed at wileyonlinelibrary.com]



(a) Distance (b) Origin (c) Destination

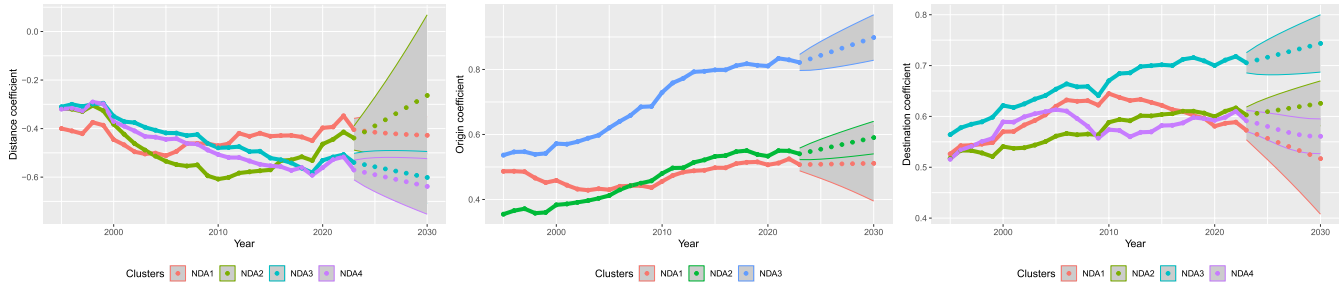
FIGURE A2 | Coefficients of gravity models for product groups in time (forecast wth ARIMA, 95% confidence interval, based on BACI-CEPII database). [Colour figure can be viewed at wileyonlinelibrary.com]

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



(a) Distance (b) Origin (c) Destination

FIGURE A3 | Coefficients of gravity models for product subgroups in time (forecast wth ARIMA, 95% confidence interval, based on BACI-CEPII database). [Colour figure can be viewed at wileyonlinelibrary.com]



(a) Distance (b) Origin (c) Destination

FIGURE A4 | Coefficients of gravity models for products in time (forecast wth ARIMA, 95% confidence interval). [Colour figure can be viewed at wileyonlinelibrary.com]

vegetables, textiles, and machinery parts; NDA_3 (approximately 200 subgroups) comprises bulk commodities, ores, and basic chemicals.

For origin GDP coefficients (Figure A3b), four clusters are identified (NDA_1 through NDA_4), with coefficients ranging from 0.20 to 0.40, reflecting varying degrees of supply-side economic scale effects across product categories.

For destination GDP coefficients (Figure A3c), three clusters show destination market size elasticities between 0.15 and 0.35, with consumer goods generally exhibiting higher demand-side GDP sensitivity than industrial intermediates.

Figure A4 shows the clustering results at the most disaggregated level—individual products (HS-6 level, more than 5000 categories)—from the CEPII-BACI database.

For distance coefficients (Figure A4a), three clusters are identified: NDA_1 exhibits the lowest distance sensitivity (around − 0.15), containing differentiated manufactured products; NDA_2 shows intermediate sensitivity (around − 0.20), comprising semi-processed goods; NDA_3 displays the highest distance sensitivity (around − 0.30), including bulk commodities and weight-sensitive products.

For origin GDP coefficients (Figure A4b), three clusters span the range from 0.25 to 0.45, with NDA_1 capturing products where exporter economic size strongly influences trade volumes, typically capital-intensive manufactures.

For destination GDP coefficients (Figure A4c), three clusters range from 0.20 to 0.30, with consumer goods and final products (NDA_1) showing higher destination GDP elasticity than intermediate inputs (NDA_2 and NDA_3).

Across all aggregation levels (Figures A2–A4), a consistent pattern emerges: the number of identified clusters ranges from two to four, with three being most common, suggesting that trade products naturally partition into distinct categories based on their sensitivity to distance and economic mass.

The correspondence between product-level clusters (BACI-CEPII) and sectoral clusters (OECD-ICIO) is examined in Section 6.6, where we discuss how product categories map to industrial sectors and whether their network-gravity relationships exhibit consistent patterns.

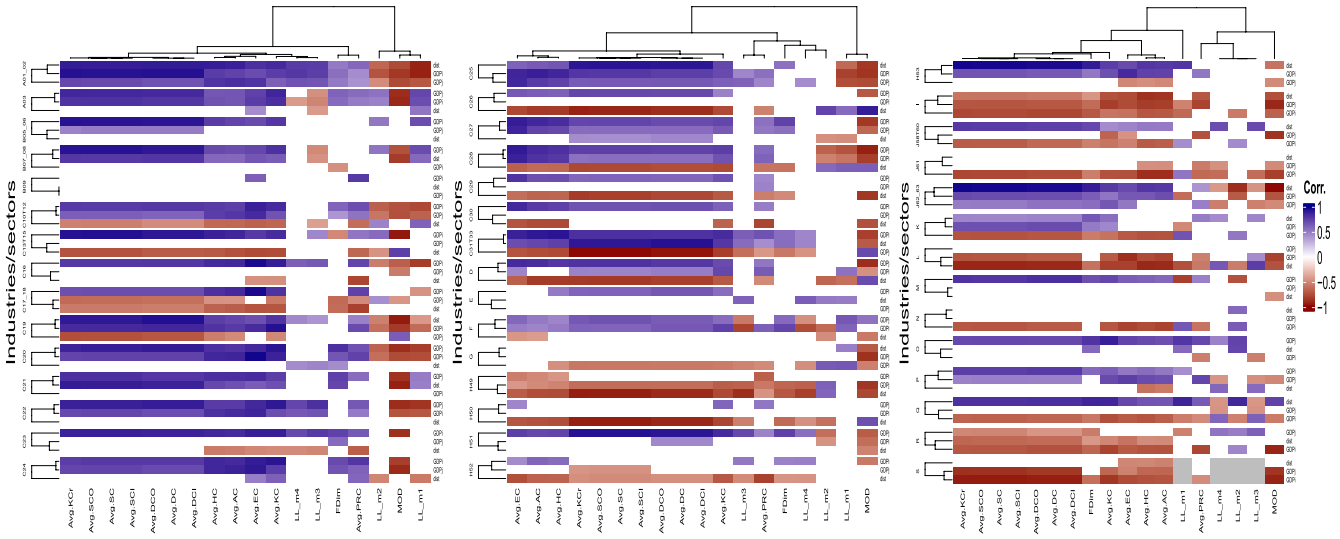
Figure A5 displays Spearman correlations between gravity coefficients and selected network indicators disaggregated by industrial sector, revealing substantial heterogeneity across industries. Cells in Figure A5 (and in Figures 5 and 6 below, which use the same convention) are set to zero and therefore appear white whenever the underlying Spearman correlation is not statistically significant at the 5% level; only coloured (non-white) cells represent correlations distinguishable from zero.

The correlation patterns vary considerably across sectors. For most manufacturing sectors (C20–C30), the ordering of correlations follows: destination GDP shows the strongest positive correlation with centrality measures, followed by origin GDP, while distance exhibits negative correlations. Service sectors (K, L, M, N) display similar patterns but with generally weaker magnitudes. Resource extraction sectors (B05_06, B07_08) show distinctive patterns where distance correlations are less negative than in manufacturing, suggesting lower distance sensitivity in raw material trade. The lacunarity and fractal dimension indicators show sector-specific patterns: technology-intensive sectors (C26, J61) exhibit stronger correlations with LL_m2 and LL_m3 than traditional manufacturing sectors.

Figure A6 presents Spearman correlations between gravity coefficients and network indicators for each of the 97 product groups in the CEPII-BACI database.

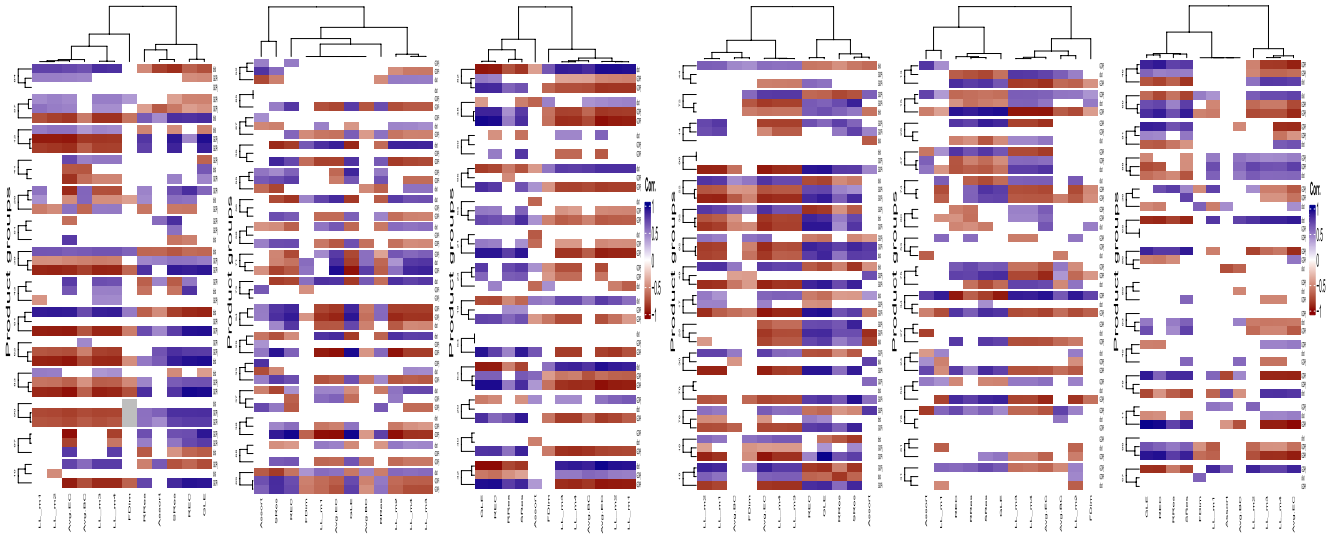
Substantial heterogeneity emerges across product groups. The lacunarity moments display strongly product-specific patterns: LL_m3 and LL_m4 correlations range from about − 0.96 to 0.92 across the 86 product groups for which the correlation could be computed, confirming that the relationship between network heterogeneity and gravity parameters depends strongly on product characteristics. Resilience indicators (RRes, SRes) are, on average, negatively—rather than positively—correlated with distance coefficients: 77% of product groups (66 of 86) show a negative RRes–distance correlation (mean $r = -0.29$) and 76% (65 of 86) show a negative SRes–distance correlation (mean $r = -0.33$), so for most product groups a more resilient network structure coincides with higher, rather than lower, distance sensitivity in trade.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



(a) Industries: A01-02–C24 (b) Industries: C25–H52 (c) Industries: H53–S

FIGURE A5 | Spearman's correlation between gravity coefficients and network-level (and aggregated node-level) indicators. [Colour figure can be viewed at wileyonlinelibrary.com]



(a) Product groups: 01–16 (b) Product groups: 17–32 (c) Product groups: 33–48 (d) Product groups: 49–64 (e) Product groups: 65–81 (f) Product groups: 83–97

FIGURE A6 | Spearman's correlation between gravity coefficients and network-level (and aggregated node-level) indicators. [Colour figure can be viewed at wileyonlinelibrary.com]

Forecast Accuracy

Table A1 shows the ARIMA and Bayesian ARIMA forecast accuracy.

Rubustness Tests of Random-Forest Causality Methods and Placebo Tests for Granger Causality

Table A2 shows the effect of shadow-feature correction.

To assess out-of-sample forecast accuracy rather than relying solely on in-sample fit, we hold out the final 5 years of each coefficient series (2016–2020 for OECD-ICIO; 2018–2022 for CEPII-BACI), re-estimate both the traditional ARIMA and Bayesian ARIMA specifications on the remaining years only, and compare the resulting forecasts to the held-out actual values using mean absolute error (MAE) and root-mean-squared error (RMSE), normalized by each series' in-sample standard deviation.

This backtest reuses the same `auto.arima()`/`auto.sarima()` model-fitting functions already used for the true out-of-sample 2023–2030 forecasts, and mirrors the blocked, expanding-window logic already applied to our Random-Forest causality tests (Section Random Forest-Based Granger-Like Causality Analysis); it introduces no new forecasting methodology.

Table A3 shows the raw versus FDR-corrected significance counts and Table A4 shows the circular-shift placebo benchmark.

Sectorial and Product Group Codes

Table A5 lists the sectorial codes.

The list of BACI-CEPII product codes.

Employed Gravity Models

Let $T_{ij,t}$ denote bilateral trade from origin country i to destination country j in period t, let $GDP_{i,t}$ and $GDP_{j,t}$ be the corresponding GDPs, and let $dist_{ij}$ be the great-circle distance between i and j (time-invariant). A parsimonious, multiplicative gravity baseline with time-varying elasticities is.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A1 | Summary of forecast accuracy.

Series	MAE	RMSE	Bayesian?
GDPj	0.0129894	0.0135896	No
dist	0.0031417	0.0041978	No
GDPi	0.0338087	0.0360924	No
GDPj	0.0071250	0.0077194	Yes
dist	0.0031006	0.0041064	Yes
GDPi	0.0405432	0.0430012	Yes
NDA(dist, sectorial)$_1$	0.0029298	0.0033131	No
NDA(dist, sectorial)$_2$	0.1000196	0.1047557	No
NDA(dist, sectorial)$_3$	0.1124182	0.1243633	No
NDA(dist, sectorial)$_1$	0.0031097	0.0033215	Yes
NDA(dist, sectorial)$_2$	0.0755771	0.0782057	Yes
NDA(dist, sectorial)$_3$	0.1137357	0.1235878	Yes
NDA(GDPi, sectorial)$_1$	0.0041857	0.0050203	No
NDA(GDPi, sectorial)$_2$	0.0024915	0.0034685	No
NDA(GDPi, sectorial)$_3$	0.0031121	0.0034378	No
NDA(GDPi, sectorial)$_4$	0.0126360	0.0156123	No
NDA(GDPi, sectorial)$_1$	0.0045678	0.0055624	Yes
NDA(GDPi, sectorial)$_2$	0.0028312	0.0038011	Yes
NDA(GDPi, sectorial)$_3$	0.0061049	0.0067377	Yes
NDA(GDPi, sectorial)$_4$	0.0354306	0.0391984	Yes
NDA(GDPj, sectorial)$_1$	0.0070790	0.0094373	No
NDA(GDPj, sectorial)$_2$	0.0460260	0.0464384	No
NDA(GDPj, sectorial)$_3$	0.0061488	0.0070157	No
NDA(GDPj, sectorial)$_1$	0.0072251	0.0095442	Yes
NDA(GDPj, sectorial)$_2$	0.0348398	0.0353784	Yes
NDA(GDPj, sectorial)$_3$	0.0274660	0.0517544	Yes
NDA(p.group, dist)$_1$	0.0065056	0.0082232	No
NDA(p.group, dist)$_2$	0.0027855	0.0035766	No
NDA(p.group, dist)$_3$	0.0303914	0.0325690	No
NDA(p.group, dist)$_1$	0.0434099	0.0522344	Yes
NDA(p.group, dist)$_2$	0.0089395	0.0114656	Yes
NDA(p.group, dist)$_3$	0.0220205	0.0236365	Yes
NDA(p.group, GDPi)$_1$	0.0324417	0.0353575	No
NDA(p.group, GDPi)$_2$	0.0157155	0.0179112	No
NDA(p.group, GDPi)$_3$	0.0555510	0.0691793	No
NDA(p.group, GDPi)$_1$	0.0285374	0.0305163	Yes
NDA(p.group, GDPi)$_2$	0.0106864	0.0138553	Yes
NDA(p.group, GDPi)$_3$	0.0377587	0.0451926	Yes
NDA(p.group, GDPj)$_1$	0.0167341	0.0187425	No
NDA(p.group, GDPj)$_2$	0.0265067	0.0300707	No
NDA(p.group, GDPj)$_3$	0.0684429	0.0758132	No
NDA(p.group, GDPj)$_4$	0.0305588	0.0341957	No
NDA(p.group, GDPj)$_1$	0.0096111	0.0105093	Yes
NDA(p.group, GDPj)$_2$	0.0072421	0.0088798	Yes
NDA(p.group, GDPj)$_3$	0.0554759	0.0606837	Yes
NDA(p.group, GDPj)$_4$	0.0245925	0.0289206	Yes

TABLE A2 | Effect of shadow-feature correction on the number of network indicators flagged as Granger-causal (Random Forest approach).

Gravity coefficient	Indicators tested	Flagged (importance > 0)	Flagged (shadow-corrected)
Distance	29	14	0
Origin GDP	29	28	23
Destination GDP	29	26	23

$$T_{ij,t} = A_t \cdot GDP_{i,t}^{\beta_{o,t}} \cdot GDP_{j,t}^{\beta_{d,t}} \cdot dist_{ij}^{-\gamma_t} \cdot \varepsilon_{ij,t}, \tag{A.1}$$

where $A_t > 0$ is a time-specific scale, $\beta_{o,t}$ and $\beta_{d,t}$ are the output elasticities at origin and destination, and γ_t is the distance elasticity ('trade-cost elasticity of distance'). In log-linear form,

$$\ln T_{ij,t} = \alpha_t + \beta_{o,t} \ln GDP_{i,t} + \beta_{d,t} \ln GDP_{j,t} - \gamma_t \ln dist_{ij} + u_{ij,t}, \tag{A.2}$$

Estimation can proceed separately for each t (sequence of cross-sections) to obtain a time series $\left\{\hat{\beta}_{o,t}, \hat{\beta}_{d,t}, \hat{\gamma}_t\right\}$, or in a panel by interacting regressors with time dummies to allow unrestricted time-variation in elasticities.

Richer specifications often augment (A.2) with additional covariates (contiguity, common language, historical ties, trade agreements, multilateral-resistance terms), improving in-sample fit and structural interpretation; see, for example, Anderson and van Wincoop (2003), Head and Mayer (2014), and Anderson (2011). However, such covariates are frequently correlated with $\ln dist_{i,j}$, and—especially in short samples—the resulting multicollinearity can change and potentially distort the estimated distance elasticity γ_t. Because our objective is to track the evolution of the core elasticities $\left\{\beta_{o,t}, \beta_{d,t}, \gamma_t\right\}$, we weigh this multicollinearity risk against the value of additional covariates when finalizing our specification. The model we actually estimate throughout this paper, Equation (1) (Section 4.1.3), extends (A.2) with nine CEPII bilateral controls $D_{m,ij}$ (Conte et al. 2022) in place of the exporter-/importer-year or country-pair fixed effects that this year-by-year cross-sectional design cannot otherwise accommodate (Section 4.1.3). Section Comparison of OLS and PPML Estimates for the Bilateral Control Variables empirically revisits the multicollinearity concern raised above: across all five databases, the nine bilateral controls show markedly weaker sign and joint-significance agreement between OLS and PPML than the three core elasticities do, consistent with their being static bilateral dummies rather than a full fixed-effects structure.

Interpretation of time-variation in coefficients. The sequence $\left\{\gamma_t\right\}$ captures how the 'penalty of distance' changes over time. Changes in the output elasticities reflect evolving size effects on bilateral trade. The resulting coefficient time series enables time series analyses such as trend estimation, structural break detection, volatility comparisons across regimes, and event studies.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A3 | Sectoral panel Granger causality (traditional method): Raw versus FDR-corrected significance counts.

Coefficient	Level	N_tested	N_sig_raw	N_sig_FDR05	N_sig_FDR10	Pct_retained_q05
Distance	Individual (sector-level)	1008	39	0	14	0.0
Origin GDP	Individual (sector-level)	1008	45	2	2	4.4
Destination GDP	Individual (sector-level)	1008	44	0	2	0.0
Distance	Pooled (panel, sector FE)	12	1	0	1	0.0
Origin GDP	Pooled (panel, sector FE)	12	9	11	11	122.2
Destination GDP	Pooled (panel, sector FE)	12	7	8	8	114.3

TABLE A4 | Circular-shift placebo benchmark ($n = 100$ replications) for the sectoral panel Granger results.

Coefficient	N_sig_observed	Placebo_mean	Placebo_SD	Placebo_p95	Empirical_p	Robust_at_5pct
Distance	117	82.1	9.2	99	0	TRUE
Origin GDP	137	75.2	7.4	86	0	TRUE
Destination GDP	130	98.3	9.5	114	0	TRUE

Disaggregation by product k, product group, or sector s reveals heterogeneity in how distance and size effects evolve across technologies, value chain positions, and demand elasticities:

$$\ln T_{ij,t}^{(k)} = \alpha_t^{(k)} + \beta_{o,t}^{(k)} \ln GDP_{i,t} + \beta_{d,t}^{(k)} \ln GDP_{j,t} - \gamma_t^{(k)} \ln dist_{ij} + u_{ij,t}^{(k)}. \tag{A.3}$$

Throughout this paper we estimate the fuller specification (1) (Section 4.1.3) by PPML on trade in levels rather than (A.2) by OLS on $\ln T_{ij,t}$, for the reasons detailed in Section A.6. Regardless of which bilateral controls are included, our object of interest remains the same three core elasticities $\{\beta_{o,t}, \beta_{d,t}, \gamma_t\}$: the 'penalty of distance' γ_t and the origin/destination output elasticities $\beta_{o,t}$, $\beta_{d,t}$.

Robustness Check

To address concerns about heteroskedasticity and zero trade flows (Santos Silva and Tenreyro 2006; Silva and Tenreyro 2011), we estimate every year-specific, product/sector-specific cross-section of Equation (1) (Section 4.1.3) by PPML rather than by log-linearized OLS on $\ln(Trade_{ij,t} + 1)$. PPML is robust to heteroskedasticity and lets zero-trade observations enter the estimation naturally, rather than being dropped or ad hoc adjusted (Santos Silva and Tenreyro 2006). As explained in Section 4.1.3, country-pair fixed effects cannot be added to this year-by-year cross-sectional design without saturating the data, and exporter- or importer-year fixed effects cannot be added without absorbing the GDP elasticities $\beta_{o,t}^k$ and $\beta_{d,t}^k$ that Equation (1) is designed to estimate; the nine CEPII bilateral controls $D_{m,ij}$ (Conte et al. 2022) take the place of either type of fixed effect.

Section Comparison of OLS and PPML Estimates for the Bilateral Control Variables reports the full empirical comparison between the OLS and PPML estimates of Equation (1) for every database, year, and product/sector, confirming that our main findings are not an artefact of the choice of estimator.

Comparison of OLS and PPML Estimates for the Bilateral Control Variables

Section A.6 explains why we estimate Equation (1) (Section 4.1.3) by PPML rather than by OLS; here we report the full empirical comparison between the two estimators. For every year t, product/sector k, and database, Equation (1) is re-estimated both by PPML on trade in levels (our baseline, reported throughout the main text) and by OLS on $\ln\left(Trade_{ij,t}^k + 1\right)$, using paired estimation scripts that share the same underlying `gravity_engine.r` routine and differ only in the link function and the treatment of zero-trade observations: `ICIO_AGG_PPML.r`/`ICIO_AGG_OLS.r` (aggregate OECD-ICIO), `ICIO_SECT_PPML.r`/`ICIO_SECT_OLS.r` (sectoral OECD-ICIO), and `CEPII_PG_PPML.r`/`CEPII_PG_OLS.r`, `CEPII_PSG_PPML.r`/`CEPII_PSG_OLS.r`, `CEPII_PROD_PPML.r`/`CEPII_PROD_OLS.r` (BACI-CEPII product groups, subgroups, and products).

For $\beta_{o,t}^k$, $\beta_{d,t}^k$, and γ_t^k we compute the Spearman rank correlation between the two series, following the procedure of Section A.6. For the nine $\delta_{m,t}^k$ bilateral-control coefficients—which, as noted in Section 4.1.3, are not carried forward as Y variables into the network-causality analysis—rather than reproducing nine additional coefficient-by-year panels in the main text, we report, for each control D_m, the share of year-by-product/sector cells in which the OLS and PPML estimates agree in sign and the share for which both are statistically significant at the 5% level.

Pooled across all five databases, years, and products/sectors, the three core elasticities agree in sign between OLS and PPML in 87.8%–98.3% of cells and are jointly significant at the 5% level under both estimators in 52.5%–91.3% of cells. The nine bilateral controls agree in sign in only 62.6%–76.9% of cells and are jointly significant under both in just 6.2%–38.3% of cells—the expected pattern for static bilateral dummies estimated without country-pair or country-year fixed effects.

Construction of Temporal Multilayer Trade Network Graphs

We model trade as a time-indexed family of weighted, directed multilayer graphs with fixed country and layer sets and time-varying edge weights.

Let $\mathcal{V} = \{1, \ldots, N\}$ denote the set of countries (nodes), $\mathcal{L} = \{1, \ldots, L\}$ the set of layers (product groups or sectors), and $\mathcal{T} = \{t_1, \ldots, t_T\}$ the set of discrete time points (years). Nodes and layers are time-invariant; only edge weights vary with $t \in \mathcal{T}$.

$$\mathcal{M} = \left\{ \mathcal{M}^t = \left(\mathcal{V}, \mathcal{L}, \mathcal{E}^t, w^t\right) \right\}_{t \in \mathcal{T}}, \tag{A.4}$$

where $\mathcal{E}^t \subseteq \mathcal{V} \times \mathcal{V} \times \mathcal{L} \times \mathcal{L}$ is the edge set at time t, and.

$$w^t : \mathcal{V} \times \mathcal{V} \times \mathcal{L} \times \mathcal{L} \to \mathbb{R}_{\geq 0}, \qquad (i,j,I,J) \mapsto w_{ij}^{I \to J}(t) \tag{A.5}$$

assigns a nonnegative trade volume from origin country i and origin layer I to destination country j and destination layer J at time t. Directedness is encoded by the ordered pair (i,j). The edge set can be written as.

$$\mathcal{E}^t = \left\{ (i,j,I,J) \in \mathcal{V}^2 \times \mathcal{L}^2 : w_{ij}^{I \to J}(t) > 0 \right\}. \tag{A.6}$$

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A5 | List of industries (sectors).

ID	Old code	Code	Industry
1	D01T02	A01_02	Agriculture, hunting, forestry
2	D03	A03	Fishing and aquaculture
3	D05T06	B05_06	Mining and quarrying, energy producing products
4	D07T08	B07_08	Mining and quarrying, non-energy producing products
5	D09	B09	Mining support service activities
6	D10T12	C10T12	Food products, beverages and tobacco
7	D13T15	C13T15	Textiles, textile products, leather and footwear
8	D16	C16	Wood and products of wood and cork
9	D17T18	C17_18	Paper products and printing
10	D19	C19	Coke and refined petroleum products
11	D20	C20	Chemical and chemical products
12	D21	C21	Pharmaceuticals, medicinal chemical and botanical products
13	D22	C22	Rubber and plastics products
14	D23	C23	Other non-metallic mineral products
15	D24	C24	Basic metals
16	D25	C25	Fabricated metal products
17	D26	C26	Computer, electronic and optical equipment
18	D27	C27	Electrical equipment
19	D28	C28	Machinery and equipment, nec
20	D29	C29	Motor vehicles, trailers and semi-trailers
21	D30	C30	Other transport equipment
22	D31T33	C31T33	Manufacturing nec; repair and installation of machinery and equipment
23	D35	D	Electricity, gas, steam and air conditioning supply
24	D36T39	E	Water supply; sewerage, waste management and remediation activities
25	D41T43	F	Construction
26	D45T47	G	Wholesale and retail trade; repair of motor vehicles
27	D49	H49	Land transport and transport via pipelines
28	D50	H50	Water transport
29	D51	H51	Air transport
30	D52	H52	Warehousing and support activities for transportation
31	D53	H53	Postal and courier activities
32	D55T56	I	Accommodation and food service activities
33	D58T60	J58T60	Publishing, audiovisual and broadcasting activities
34	D61	J61	Telecommunications
35	D62T63	J62_63	IT and other information services
36	D64T66	K	Financial and insurance activities
37	D68	L	Real estate activities
38	D69T75	M	Professional, scientific and technical activities
39	D77T82	N	Administrative and support services
40	D84	O	Public administration and defence; compulsory social security
41	D85	P	Education
42	D86T88	Q	Human health and social work activities
43	D90T93	R	Arts, entertainment and recreation
44	D94T96	S	Other service activities
45	D97T98	T	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use

$$A(t)_{ijIJ} = w_{ij}^{I\to J}(t), \tag{A.7}$$

that is, $A(t) \in \mathbb{R}_{\geq 0}^{N\times N\times L\times L}$. Equivalently, the $NL \times NL$ supra-adjacency matrix $S(t)$ indexed by pairs (i, I) and (j, J) is.

$$S(t)_{(i,I),(j,J)} = w_{ij}^{I\to J}(t), \qquad S(t) \in \mathbb{R}_{\geq 0}^{NL\times NL}. \tag{A.8}$$

Writing $S(t)$ in $L \times L$ blocks of size $N \times N$ yields.

$$S(t) = \begin{bmatrix} B_{11}(t) & B_{12}(t) & \cdots & B_{1L}(t) \\ B_{21}(t) & B_{22}(t) & \cdots & B_{2L}(t) \\ \vdots & \vdots & \ddots & \vdots \\ B_{L1}(t) & B_{L2}(t) & \cdots & B_{LL}(t) \end{bmatrix}, \qquad B_{IJ}(t) = \left[w_{ij}^{I\to J}(t)\right]_{i,j=1}^{N}. \tag{A.9}$$

The temporal network is the sequence with fixed dimensions and time-varying entries:

$$\{S(t)\}_{t\in\mathcal{T}}, \qquad N, L \text{ fixed}, S(t) \text{ varies over } t. \tag{A.10}$$

No inter-time edges are assumed by default; dynamic processes can be modelled by introducing inter-slice coupling between (i, I, t) and $(i, I, t + 1)$ if required.

Calculated Network Properties

Table A7 lists of employed node-level indicators.

Table A8 lists the employed similarity indicators.

Table A9 lists the employed network-level indicators.

Table A10 shows the connection between network properties and the coefficients of gravity models.

Gravity Model Coefficients Context:

Distance dependence (β_{dist}): Reflected in AVPL, GLE, ALE, ECC, SP—network measures of separation and efficiency capture the role of geographic/economic distance.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A6 | BACI-CEPII Database product classes up to 2 digits.

HS code	Product class
01–24	Live animals; animal products
25–27	Vegetable products
28–38	Foodstuffs, beverages, and tobacco
39–40	Mineral products
41–43	Raw hides, leather, furs, skins, and articles thereof
44–49	Wood and articles of wood; wood charcoal; cork and articles of cork; articles of straw, of esparto or of other plaiting materials; basketware and wickerwork
50–63	Textiles and textile articles
64–67	Footwear, headgear, and parts thereof
68–71	Stone, plaster, cement, and articles thereof; ceramic products; glass and glassware
72–83	Metals and articles thereof
84–85	Machinery and mechanical appliances; electrical equipment; parts thereof; sound recorders and reproducers, television image and sound recorders and reproducers, and parts and accessories of such articles
86–89	Vehicles, aircraft, vessels, and associated transport equipment
90–92	Optical, photographic, cinematographic, measuring, checking, precision, medical or surgical instruments and apparatus; clocks and watches; musical instruments; parts and accessories thereof
93–96	Arms and ammunition; parts and accessories thereof
97–99	Miscellaneous goods, including works of art, collectors' items, and antiques

Economic mass—Exporter (β_{GDP_i}): Reflected in SCO, DCO, HC, PRC—measures of export capacity and influence capture sender country economic role.

Economic mass—Importer (β_{GDP_j}): Reflected in SCI, DCI, AC—measures of import capacity and authority capture receiver country economic role.

Structural heterogeneity: FDim, LL moments—fractal and lacunarity measures capture scale-free and hierarchical properties that may modulate gravity coefficients over time. Whereas reflected in MOD, Assortativity, ES, NS—community structure and similarity measures capture institutional and political economy factors.

Research Question Context: Time-varying gravity coefficients may be driven by or correlated with changes in network topology. For instance:

- Increasing LL_m1 or LL_m2 (greater heterogeneity) may correlate with changing β_{dist} as countries adopt diverse strategies.
- Rising MOD may correspond to decreasing FTA coefficients within blocs and increasing border effects between blocs.
- Changes in FDim may signal structural phase transitions affecting all gravity coefficients.
- Granger causality tests can establish whether network indicators predict changes in gravity coefficients or vice versa, addressing endogeneity.

Pattern Recognition With Network-Based Dimensionality Analysis

The employed GNDA (Kosztyán et al. 2024) provides a robust framework for clustering high-dimensional time series data, particularly suited for scenarios with limited observations but numerous variables—characteristic of retail network data with multiple products and sectors.

GNDA addresses the fundamental challenge of determining the optimal number of Latent Variables (LVs) in High Dimensional Low Sample Size (HDLSS) datasets. Unlike traditional methods such as Principal Component Analysis (PCA) or Principal Factor Analysis (PFA), which often over- or underestimate the number of factors, GNDA leverages network science and modularity-based community detection to identify natural groupings of time series characteristics.

The methodology transforms the correlation structure of variables into a network representation. Let $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_m]$ denote the standardized time series matrix with n observations and m variables, where each $\mathbf{z}_i$ represents a standardized time series:

$$\mathbf{z}_i = \frac{\mathbf{v}_i - \mu_{\mathbf{v}_i}}{\sigma_{\mathbf{v}_i}} \tag{A.11}$$

The similarity graph $G(\mathcal{V}, \mathcal{E}, \mathcal{W})$ is constructed where nodes $i \in \mathcal{V}$ represent variables, edges $(i,j) \in \mathcal{E}$ encode relationships, and weights $w_{i,j} \in \mathcal{W}$ quantify similarity. For time series clustering, the squared correlation $r_{i,j}^2$ serves as the similarity measure, though distance correlation or other metrics may be employed depending on the data structure.

Community detection partitions the graph into N modules $C_1, C_2, \ldots, C_N$ by maximizing the modularity function:

$$M = \frac{1}{2L} \sum_{i,j} \left(r_{i,j} - \gamma \hat{r}_{i,j}\right) \delta\left(C_i, C_j\right) \tag{A.12}$$

where L is the total edge weight, $\hat{r}_{i,j}$ represents the expected weight under a null model, γ is the resolution parameter, and $\delta\left(C_i, C_j\right) = 1$ if nodes i and j belong to the same community, zero otherwise.

For each identified module C_I, the corresponding latent variable is constructed as a weighted combination of standardized time series using EC values as weights:

$$LV_I = \frac{\sum_{i \in C_I} c_i \mathbf{z}_i}{\sum_{i \in C_I} c_i} \tag{A.13}$$

where c_i denotes the eigenvector centrality of node i. The EC is defined through the dominant eigenvector of the adjacency matrix $\mathbf{A}$:

$$c_k = \frac{1}{\lambda} \sum_{l \in G} a_{k,l} c_l \tag{A.14}$$

where λ is the largest eigenvalue. This centrality measure provides a natural interpretation analogous to factor loadings in traditional factor analysis, with higher values indicating greater importance within the cluster.

The critical advantage of GNDA for time series clustering lies in its automated determination of the number of clusters through modularity optimization, eliminating the subjective specification required by most dimensionality reduction techniques. The method performs robustly even when $n \ll m$, as demonstrated on simulated block matrices where GNDA consistently identified the correct number of latent variables while PCA and PFA frequently failed.

The resulting latent time series LV_I represents the characteristic temporal pattern of cluster I, capturing the common dynamics shared by its constituent series. These cluster-level time series can be directly utilized for forecasting, providing aggregate predictions that reflect the collective behaviour of related variables. The EVC-weighted aggregation ensures that more central (representative) series contribute proportionally more to the latent variable, yielding a robust characterization of cluster dynamics.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A7 | List of employed node-level indicators.

Abbreviations	Indicator	Calculation formula	Brief interpretation
DCI	In-degree Centrality	$\mathrm{DCI}_i = \sum_{j=1}^{N} a_{ji}$	Number of incoming trade connections; measures import partner diversity
DCO	Out-degree Centrality	$\mathrm{DCO}_i = \sum_{j=1}^{N} a_{ij}$	Number of outgoing trade connections; measures export partner diversity
SCI	In-strength Centrality	$\mathrm{SCI}_i = \sum_{j=1}^{N} w_{ji}$	Total value of imports; captures import volume from all partners
SCO	Out-strength Centrality	$\mathrm{SCO}_i = \sum_{j=1}^{N} w_{ij}$	Total value of exports; captures export volume to all partners
BC	Betweenness Centrality	$\mathrm{BC}_i = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}$	Fraction of shortest trade paths passing through node i; measures brokerage role
EC	Eigenvector Centrality	$\mathbf{Ax} = \lambda_{\max}\mathbf{x}$, where $\mathrm{EC}_i = x_i$	Centrality based on connections to other central nodes; captures prestige
KC	Katz Centrality	$\mathrm{KC}_i = \alpha \sum_{j=1}^{N} a_{ij}\mathrm{KC}_j + \beta$	Weighted sum of direct and indirect connections; accounts for all path lengths
HC	Hub Centrality	$\mathbf{h} = \mathbf{Aa}$; $\mathbf{a} = \mathbf{A}^T\mathbf{h}$ (iterative)	Authority of nodes pointed to; identifies countries exporting to important importers
AC	Authority Centrality	$\mathbf{a} = \mathbf{A}^T\mathbf{h}$; $\mathbf{h} = \mathbf{Aa}$ (iterative)	Authority of nodes pointing in; identifies countries importing from important exporters
PRC	PageRank Centrality	$\mathrm{PRC}_i = (1-d) + d \sum_{j \in \mathcal{N}_i^{\mathrm{in}}} \frac{\mathrm{PRC}_j}{\mathrm{DCO}_j}$	Recursive prestige measure; importance weighted by partners' importance
KCR	K-core	$\max\{k: \mathrm{DCI}_i \geq k \text{ and } \mathrm{DCO}_i \geq k\}$	Maximal subgraph where all nodes have degree $\geq k$; identifies core-periphery structure
ECC	Eccentricity	$\mathrm{ECC}_i = \max_j d(i,j)$	Maximum shortest path distance from node i; measures remoteness
LEV	Local Efficiency Value	$\mathrm{LEV}_i = \frac{1}{\lvert\mathcal{N}_i\rvert(\lvert\mathcal{N}_i\rvert-1)} \sum_{j,k \in \mathcal{N}_i, j \neq k} \frac{1}{d_{jk}}$	Efficiency of subgraph of node i's neighbours; measures local connectivity robustness
REC	Reciprocity	$REC = \frac{\sum_{i \neq j} a_{ij} \cdot a_{ji}}{\sum_{i \neq j} a_{ij}}$	The tendency of pairs of nodes to form mutual connections. It is defined as the proportion of links that are bidirectional relative to the total number of links in the network
LL	Local Lacunarity	$\mathrm{LL}_i = \frac{\langle r_i^2 \rangle}{\langle r_i \rangle^2}$	Local heterogeneity of trade connections; measures irregularity around node i

Note: N = number of nodes (countries); a_{ij} = adjacency matrix element (1 if edge from i to j, 0 otherwise); w_{ij} = weighted adjacency matrix element (trade flow value); $\mathbf{A}$ = adjacency matrix; σ_{st} = number of shortest paths between nodes s and t; $\sigma_{st}(i)$ = number of those paths passing through i; $\lambda_{\max}$ = largest eigenvalue of $\mathbf{A}$; $\mathbf{x}$ = eigenvector corresponding to $\lambda_{\max}$; α, β = Katz parameters; d = damping factor (typically 0.85); $\mathcal{N}_i$ = neighbourhood of node i; $\mathcal{N}_i^{\mathrm{in}}$ = in-neighbours of node i; $d(i,j)$ = shortest path distance between nodes i and j; d_{jk} = shortest path distance in subgraph; r_i = local mass distribution within box around node i; $\langle \cdot \rangle$ = ensemble average.

TABLE A8 | List of employed layer similarity indicators.

Abbreviations	Indicator	Calculation formula	Brief interpretation
NS	Node Similarity (Overlap)	$\mathrm{NS}(t_1,t_2) = \frac{\lvert V_{t_1} \cap V_{t_2}\rvert}{\lvert V_{t_1} \cup V_{t_2}\rvert}$	Fraction of countries active in both time periods/ layers; measures stability
ES	Edge Similarity (Overlap)	$\mathrm{ES}(t_1,t_2) = \frac{\lvert E_{t_1} \cap E_{t_2}\rvert}{\lvert E_{t_1} \cup E_{t_2}\rvert}$	Fraction of trade links present in both layers; measures relationship persistence
PD	Degree-Degree Pearson Correlation	$\mathrm{PD}(t_1,t_2) = \frac{\mathrm{cov}(\mathbf{k}_{t_1},\mathbf{k}_{t_2})}{\sigma_{k_{t_1}}\sigma_{k_{t_2}}}$	Linear correlation of node degrees across layers; assesses structural similarity
SD	Degree-Degree Spearman Correlation	$\mathrm{SD}(t_1,t_2) = 1 - \frac{6\sum_i (R_{i,t_1} - R_{i,t_2})^2}{n(n^2-1)}$	Rank correlation of node degrees across layers; robust to outliers
SP	Shortest Path Similarity	$\mathrm{SP}(t_1,t_2) = 1 - \frac{1}{N(N-1)} \sum_{i \neq j} \frac{\lvert d_{ij}^{t_1} - d_{ij}^{t_2}\rvert}{\max(d_{ij}^{t_1}, d_{ij}^{t_2})}$	Normalized distance between shortest path matrices; measures topological evolution

Note: t_1, t_2 = two time periods or network layers; V_t = set of nodes (countries) in layer t; E_t = set of edges (trade links) in layer t; $\lvert\cdot\rvert$ = cardinality (size) of set; $\mathbf{k}_t$ = vector of node degrees in layer t; $\mathrm{cov}(\cdot,\cdot)$ = covariance; σ_{k_t} = standard deviation of degrees in layer t; $R_{i,t}$ = rank of node i's degree in layer t; n = number of common nodes; d_{ij}^t = shortest path distance between nodes i and j in layer t.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A9 | List of employed network-level indicators.

Abbreviations	Indicator	Calculation formula	Brief interpretation
Dens	Density	$\text{Dens} = \frac{\lvert E \rvert}{N(N-1)}$	Proportion of actual to possible trade links; measures network completeness
Assort	Assortativity	$\text{Assort} = \frac{\sum_{ij}\left(a_{ij} - k_i k_j / 2\lvert E \rvert\right)\left(k_i - \bar{k}\right)\left(k_j - \bar{k}\right)}{\sum_i k_i \left(k_i - \bar{k}\right)^2}$	Tendency of similar-degree nodes to connect; positive indicates homophily
Trace	Trace	$\sum_{i=1}^{N} a_{ii}$	Represents the total volume of self-loops. It captures the extent of domestic or intra-industial transactions within the network
DZI	In-Degree Centralization	$\text{DZI} = \frac{\sum_{i=1}^{N}\left(\text{DCI}_{\max} - \text{DCI}_i\right)}{(N-1)^2}$	Variation in in-degrees; high values indicate import concentration
DZO	Out-Degree Centralization	$\text{DZO} = \frac{\sum_{i=1}^{N}\left(\text{DCO}_{\max} - \text{DCO}_i\right)}{(N-1)^2}$	Variation in out-degrees; high values indicate export concentration
BZ	Betweenness Centralization	$\text{BZ} = \frac{\sum_{i=1}^{N}\left(\text{BC}_{\max} - \text{BC}_i\right)}{(N-1)(N-2)/2}$	Variation in betweenness; high values indicate reliance on key intermediaries
PZ	Power Centralization	$\text{PZ} = \frac{\sum_{i=1}^{N}\left(\text{EC}_{\max} - \text{EC}_i\right)}{\max\left(\sum_{i=1}^{N}\left(\text{EC}_{\max} - \text{EC}_i\right)\right)}$	Variation in eigenvector centrality; measures dominance hierarchy
HZ	Harmonic Centralization	$\text{HZ} = \frac{\sum_{i=1}^{N}\left(H_{\max} - H_i\right)}{(N-1)(N-2)}$ where $H_i = \sum_{j \neq i} \frac{1}{d(i,j)}$	Variation in harmonic centrality; assesses accessibility inequality
PRZ	PageRank Centralization	$\text{PRZ} = \frac{\sum_{i=1}^{N}\left(\text{PRC}_{\max} - \text{PRC}_i\right)}{\frac{N-1}{N}}$	Variation in PageRank; measures influence concentration
MOD	(Leiden's) Modularity Value	$\text{MOD} = \frac{1}{2\lvert E \rvert}\sum_{ij}\left(a_{ij} - \frac{k_i k_j}{2\lvert E \rvert}\right)\delta\left(c_i, c_j\right)$	Strength of community structure; high values indicate regional trade blocs
Trans	Transitivity	$\text{Trans} = \frac{3 \times \text{number of triangles}}{\text{number of connected triples}}$	Global clustering coefficient; measures triplet closure tendency
GLE	Global Efficiency	$\text{GLE} = \frac{1}{N(N-1)}\sum_{i \neq j}\frac{1}{d(i,j)}$	Average inverse shortest path length; measures network-wide connectivity
ALE	Average Local Efficiency	$\text{ALE} = \frac{1}{N}\sum_{i=1}^{N}\text{LEV}_i$	Mean of local efficiencies; measures robustness to local disruptions
AVPL	Average Path Length	$\text{AVPL} = \frac{1}{N(N-1)}\sum_{i \neq j} d(i,j)$	Mean shortest path distance; measures typical separation between countries
MOT	3-Motifs	$\text{MOT}_m = \frac{n_m}{\sum_{m'} n_{m'}}$	Frequency of 3-node subgraph pattern m; reveals local structural motifs
RRes	Resilience (Random Attack)	$\text{RRes} = \frac{1}{N}\sum_{i=1}^{N}\frac{S_i}{N}$	Area under size-of-largest-component curve during random removal
SRes	Resilience (Systematic Attack)	$\text{SRes} = \frac{1}{N}\sum_{i=1}^{N}\frac{S_i^*}{N}$	Area under curve during targeted removal (by centrality); assesses vulnerability
FDim	Fractal Dimension	$\text{FDim} = -\frac{d\ln N_B(l_B)}{d\ln l_B}$	Self-similarity exponent from box-covering; characterizes scale-free structure
LL_m1	Mean of Local Lacunarity	$\text{LL_m1} = \mu_{LL} = \frac{1}{N}\sum_{i=1}^{N}\text{LL}_i$	Average heterogeneity of local structures; indicates overall irregularity

(Continues)

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A9 | (Continued)

Abbreviations	Indicator	Calculation formula	Brief interpretation
LL_m2	Stdandard Deviation of Local Lacunarity	$\text{LL_m2} = \sigma_{LL} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\text{LL}_i - \mu_{LL}^2\right)}$	Dispersion of local heterogeneities; measures structural diversity
LL_m3	Skewness of Local Lacunarity	$\text{LL_m3} = \gamma_{LL} = \frac{1}{N\sigma_{LL}^3}\sum_{i=1}^{N}\left(\text{LL}_i - \mu_{LL}\right)^3$	Asymmetry of lacunarity distribution; detects concentration patterns
LL_m4	Kurtosis of Local Lacunarity	$\text{LL_m4} = \kappa_{LL} = \frac{1}{N\sigma_{LL}^4}\sum_{i=1}^{N}\left(\text{LL}_i - \mu_{LL}\right)^4$	Tail heaviness of lacunarity distribution; identifies extreme heterogeneity

Note: $|E|$ = number of edges; k_i = degree of node i; $\bar{k}$ = average degree; $\text{DCI}_{\max}, \text{DCO}_{\max}, \text{BC}_{\max}, \text{EC}_{\max}, \text{PRC}_{\max}$ = maximum values of respective centralities; H_i = harmonic centrality of node i; c_i = community assignment of node i; $\delta(c_i, c_j)$ = Kronecker delta (1 if $c_i = c_j$, 0 otherwise); $E_{>k}$ = number of edges among nodes with degree > k; $N_{>k}$ = number of nodes with degree > k; n_m = count of motif m; S_i = size of largest connected component after removing i nodes; S_i^* = size during targeted removal; $N_B(l_B)$ = number of boxes of size l_B needed to cover network.

Forecastings

Traditional and ARIMA and Bayesian ARIMA Models

Let $y_t \in \mathbb{R}$ denote a univariate time series observed for $t = 1, \ldots, T$. With the lag operator L, $Ly_t = y_{t-1}$, an ARIMA(p,d,q) model can be written in compact lag-polynomial form as.

$$\phi(L)(1-L)^d y_t = c + \theta(L)\varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2), \tag{A.15}$$

where $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$ and $\theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q$ enforce, respectively, stationarity (roots of ϕ outside the unit circle) and invertibility (roots of θ outside the unit circle). The operator $(1-L)^d$ encodes d-fold differencing. In levels, (A.15) may be equivalently written as.

$$(1-L)^d y_t = \mu + \sum_{i=1}^{p} \phi_i (1-L)^d y_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}, \tag{A.16}$$

with an appropriate intercept μ (linked to c once differencing is expanded).

Maximum-Likehood Estimation (MLE) under Gaussian innovations proceeds by maximizing the exact Gaussian likelihood of the innovation sequence $\{\varepsilon_t(\theta)\}$ implied by (A.15):

$$\ell(\theta; y_{1:T}) = -\frac{T}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\varepsilon_t(\theta)^2, \tag{A.17}$$

where $\theta = (\mu, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, \sigma^2)$ and the innovations are computed via the innovations algorithm or an equivalent state-space (Kalman filter) representation. Under regularity, $\widehat{\theta}_{\text{MLE}}$ is consistent and asymptotically normal, but for short samples the sampling variability of $\widehat{\theta}_{\text{MLE}}$ may be substantial, and model-order uncertainty (p, d, q) can materially affect inference and prediction.

Bayesian ARIMA. A Bayesian specification places a prior density $\pi(\theta)$ over the parameter vector θ, truncated to the stationarity/invertibility region. The posterior is.

$$p(\theta|y_{1:T}) \propto L(\theta; y_{1:T})\,\pi(\theta), \tag{A.18}$$

where $L(\theta; y_{1:T}) = \exp\{\ell(\theta; y_{1:T})\}$ is the Gaussian likelihood (A.17). Typical weakly informative choices include Gaussian priors on $(\mu, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)$ (with transforms enforcing the root constraints, e.g., via partial autocorrelations) and inverse-gamma or half-t/half-Cauchy priors for σ^2. Posterior inference is obtained by Markov Chain Monte Carlo (MCMC) (e.g., Metropolis–Hastings, Gibbs-within-Metropolis, or HMC/NUTS).

A key advantage for short time series is that Bayesian inference integrates parameter and model-order uncertainty instead of conditioning on a single point estimate or order. The h-step posterior predictive distribution is.

$$p(y_{T+h}|y_{1:T}) = \int p(y_{T+h}|\theta, y_{1:T})\,p(\theta|y_{1:T})\,d\theta, \tag{A.19}$$

which averages forecasts over the posterior of θ; numerically, (A.19) is approximated by Monte Carlo averages of conditional ARIMA forecasts across MCMC draws of θ. This automatically widens (and often better calibrates) predictive intervals when T is small.

Conjugate special case (AR regression). When $q = 0$ and $d = 0$, AR(p) can be treated as a Gaussian regression.

$$y_t = \mu + \phi^\top y_{t-1:t-p} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2), \tag{A.20}$$

with $y_{t-1:t-p} = (y_{t-1}, \ldots, y_{t-p})^\top$. Using a Normal–Inverse-Gamma prior.

$$(\phi, \mu) \mid \sigma^2 \sim \mathcal{N}(m_0, \sigma^2 V_0), \qquad \sigma^2 \sim \text{IG}(a_0, b_0), \tag{A.21}$$

yields closed-form posteriors.

$$V_n = (V_0^{-1} + X^\top X)^{-1}, \qquad m_n = V_n(V_0^{-1} m_0 + X^\top y), \tag{A.22}$$

$$a_n = a_0 + \frac{n}{2}, \qquad b_n = b_0 + \frac{1}{2}\left(y^\top y + m_0^\top V_0^{-1} m_0 - m_n^\top V_n^{-1} m_n\right), \tag{A.23}$$

where X collects the lag regressors and intercept, and n is the effective sample. This highlights Bayesian shrinkage: with short n, (m_0, V_0) regularize (ϕ, μ), reducing variance and improving predictive stability.

Model comparison and order uncertainty. Bayesian model comparison across candidate orders (p, d, q) uses Bayes factors or posterior model probabilities. The Bayes factor for model $\mathcal{M}_1$ against $\mathcal{M}_0$ is.

$$\text{BF}_{10} = \frac{m(y_{1:T}|\mathcal{M}_1)}{m(y_{1:T}|\mathcal{M}_0)} = \frac{\int L(\theta_1; y)\,\pi(\theta_1)\,d\theta_1}{\int L(\theta_0; y)\,\pi(\theta_0)\,d\theta_0}. \tag{A.24}$$

When computing $m(\cdot)$ exactly is costly, the Schwarz approximation gives.

$$\begin{aligned} \log \text{BF}_{10} &\approx -1/2\left(\text{BIC}(\mathcal{M}_1) - \text{BIC}(\mathcal{M}_0)\right), \\ \text{BIC}(\mathcal{M}) &= -2\ell\left(\widehat{\theta}_{\mathcal{M}}\right) + k_{\mathcal{M}} \log T, \end{aligned} \tag{A.25}$$

which is particularly convenient in short samples (stable and fast) and coheres with classical information criteria.

Why Bayesian methods help with short series. For small T, (i) priors induce shrinkage that mitigates overfitting of ϕ and θ; (ii) posterior predictive (A.19) accounts for parameter uncertainty that MLE-based plug-in forecasts ignore; (iii) model uncertainty over (p, d, q) can be integrated rather than decided by a single pretest; and (iv) stationarity/invertibility constraints are enforced by the prior support or parameter transforms. Collectively, these features improve finite-sample calibration of

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE A10 | Connection between network indicators and the coefficients of gravity models.

Indicator	Gravity coefficient link	Interpretation in trade context	Reflected phenomena
Node-level indicators			
DCI, DCO	Border effects, bilateral agreements	Partner diversification; export/import market breadth	Trade liberalization, regionalization
SCI, SCO	Economic mass (GDP) coefficients	Total trade volume; economic size manifestation	Growth, market size evolution
BC	Transit trade, intermediation	Brokerage in global value chains	Hub emergence, re-export centers
EC, PRC, KC	Prestige, preferential attachment	Trading with important partners amplifies centrality	Core-periphery formation
HC, AC	Directional influence	Export-to-importers vs. import-from-exporters patterns	Value chain positioning
KCr	Core-periphery structure	Integration level in dense trading core	Marginalization vs. integration
Eccentricity (ECC), Local Efficiency Value (LEV)	Remoteness, integration	Accessibility; vulnerability to local disruptions	Geographic/economic isolation
LL	Heterogeneity in connections	Irregularity of trade partner distribution	Specialized vs. diversified strategies
Similarity indicators			
Node Similarity (NS), Edge Similarity (ES)	Structural stability	Persistence of countries and links over time	Entry/exit dynamics, relationship longevity
Degree-Degree Pearson Correlation (PD), Degree-Degree Spearman Correlation (SD)	Rank stability	Consistency of trade hierarchy across periods	Structural breaks, regime shifts
Shortest Path Similarity (SP)	Topological evolution	Changes in indirect connectivity patterns	Network reorganization, path dependencies
Network-level indicators			
Density	Overall trade intensity	Proportion of realized trade relationships	Globalization vs. fragmentation
Assortativity	Preferential attachment	Rich-trade-with-rich or rich-trade-with-poor	Inequality, development convergence
DZI, DZO	Market concentration	Dominance of few import/export hubs	Dependence, systemic risk
BZ, PZ, HZ, PRZ	Hierarchy, influence distribution	Centralization of control and influence	Power imbalances, hegemony
MOD	Regional trade agreements	Strength of trade bloc formation	Regionalization, FTA impacts
Transitivity	Triadic closure	Tendency for partners of partners to trade	Trust networks, referral effects
GLE, ALE	Distance decay	Overall and local connectivity efficiency	Transport costs, infrastructure quality
AVPL	Distance coefficient	Typical separation in trade network	Gravity distance elasticity manifestation
Rich-Club Coefficient (RCC)	Elite club formation	Preferential connectivity among large economies	G7/G20 dominance
MOT	Structural patterns	Recurring local configurations (mutual, chain)	Specific trade arrangements
RRes, SRes	Shock propagation	Robustness to random vs. targeted disruptions	Crisis vulnerability, contagion
FDim	Scale-free properties	Self-similarity across scales	Universal growth mechanisms
LL_m1, LL_m2	Average heterogeneity	Overall and variability of structural irregularity	Diversity of trade strategies
LL_m3, LL_m4	Distribution shape	Skewness and tail behaviour of heterogeneity	Presence of outliers, extreme specialization

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

uncertainty and reduce spurious dynamics—properties that are crucial when only short annual histories are available.

Causality Analysis

Implementation and Imputation

All three methods require complete time series. We handle missing values via linear interpolation using the `imputeTS` package (Moritz and Bartz-Beielstein 2017). For each sector j and variable i, gaps in $Y_{j,t}$ and $X_{i,j,t}$ are filled by linear interpolation between observed neighbours, preserving temporal structure.

Each panel Granger causal test method produces an $M \times (J+1)$ matrix $\mathbf{P} = [\hat{p}_{i,j}]$, where rows correspond to explanatory variables, columns $1, \ldots, J$ to individual sectors, and column $J+1$ to the pooled panel. An entry $\hat{p}_{i,j} > 0$ indicates that variable i Granger-causes sector j's outcome at lag $\hat{p}_{i,j}$; an entry of zero indicates no detectable causality. Comparing matrices across methods reveals robustness: relationships detected by all three approaches are likely genuine, whereas method-specific findings may reflect linearity assumptions (traditional and Bayesian) or nonlinear patterns (Random Forest).

The lag estimates $\{\hat{p}_{i,j}\}$ inform the temporal dynamics of causal relationships, enabling time series analyses such as impulse-response functions, forecasting models, and event studies that respect the identified causal lags. Sector-level heterogeneity in $\hat{p}_{i,j}$ reveals differential adjustment speeds or transmission mechanisms across sectors, analogous to the heterogeneity in distance elasticities $\gamma_t^{(k)}$ discussed in the gravity model context.

Traditional and Bayesian Granger Causality Analysis

Let $y_t \in \mathbb{R}$ denote the scalar target (explained) series and $X_t = (x_{1,t}, \ldots, x_{d,t})^\top \in \mathbb{R}^d$ the vector of candidate explanatory (predictor) series observed for $t = 1, \ldots, T$. Fix a maximum lag order $p \geq 1$. lag-ℓ values are written as $y_{t-\ell}$ and $x_{j,t-\ell}$, and index sets over variables and lags are finite.

We use two closely related linear formulations. The first is a full ARX(p) model for y_t involving its own lags and all X-lags:

$$y_t = c + \sum_{\ell=1}^{p} \phi_\ell y_{t-\ell} + \sum_{j=1}^{d} \sum_{\ell=1}^{p} \beta_{j,\ell} x_{j,t-\ell} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2), \tag{A.26}$$

and the second is a bivariate *VAR*(p) for each pair (y, x_j), written with $Y_{j,t} = (y_t, x_{j,t})^\top$,

$$Y_{j,t} = a_j + \sum_{\ell=1}^{p} A_{j,\ell} Y_{j,t-\ell} + u_{j,t}, \qquad u_{j,t} \sim \mathcal{N}(0, \Sigma_j). \tag{A.27}$$

Classical Granger causality of $x_j \to y$ asks whether past x_j helps predict y beyond y's own past and (when applicable) the past of the remaining variables in the model.

In the ARX formulation (A.26), Granger non-causality of x_j with respect to y is the linear restriction.

$$H_{0,j}: \beta_{j,1} = \cdots = \beta_{j,p} = 0 \text{ vs. } H_{1,j}: \exists \ell \in \{1, \ldots, p\} \text{ such that } \beta_{j,\ell} \neq 0. \tag{A.28}$$

In the bivariate VAR formulation (A.27), Granger non-causality corresponds to zeroing all coefficients of lagged x_j in the y-equation of $Y_{j,t}$.

We use three complementary evidential tools: (i) a classical nested F–test; (ii) an information-criterion check via BIC; and (iii) Bayesian validation combining Bayes factors and MCMC-based credible intervals. A causal link is accepted only when the pre-specified classical significance and the Bayesian conditions jointly hold; the information criterion serves as an additional strength-of-evidence measure and as a basis for importance metrics.

Classical nested F–test (ARX). Let $\mathcal{M}_{\text{full}}$ be (A.26) and $\mathcal{M}_{\text{red},j}$ the reduced model obtained by removing the p regressors $\{x_{j,t-\ell}\}_{\ell=1}^{p}$. Denote by RSS_{full} and $\text{RSS}_{\text{red},j}$ the residual sum of squares, q_j the number of restrictions ($= p$), and k_{full} the number of free parameters in $\mathcal{M}_{\text{full}}$. The standard Wald/F statistic is.

$$F_j = \frac{(\text{RSS}_{\text{red},j} - \text{RSS}_{\text{full}})/q_j}{\text{RSS}_{\text{full}}/(T - k_{\text{full}})} \sim F(q_j, T - k_{\text{full}}) \text{ under } H_{0,j}. \tag{A.29}$$

We reject $H_{0,j}$ at level α when the p–value p_j satisfies $p_j \leq \alpha$. In the bivariate VAR test, the same logic applies but the restriction is imposed only on the y–equation of (A.27).

Information-criterion (BIC) check. As an information-theoretic control that is less sensitive to finite-sample variance, we also compute the BIC difference between the reduced and full models:

$$\Delta\text{BIC}_j = \text{BIC}(\mathcal{M}_{\text{red},j}) - \text{BIC}(\mathcal{M}_{\text{full}}), \ \text{BIC}(\mathcal{M}) = -2 \log L\left(\hat{\theta}_{\mathcal{M}}\right) + k_{\mathcal{M}} \log T. \tag{A.30}$$

A positive ΔBIC_j indicates that including the lags of x_j improves the information criterion and therefore supports $x_j \to y$. In practice we use ΔBIC_j primarily as (i) an evidence-strength score consistent with the classical test; and (ii) a building block for relative-importance metrics.

Expected delay (lag selection) and lag-level importance. For each predictor x_j we estimate the 'dominant' delay by the lag index achieving the largest absolute t–statistic in the unrestricted y–equation:

$$\hat{\ell}_j = \arg\max_{\ell \in \{1,\ldots,p\}} |t_{j,\ell}|, \ t_{j,\ell} = \frac{\hat{\beta}_{j,\ell}}{\widehat{\text{se}}\left(\hat{\beta}_{j,\ell}\right)}. \tag{A.31}$$

If $H_{0,j}$ is not rejected (classically and Bayes-wise; see below), we set $\hat{\ell}_j = 0$. Within-variable lag importance profiles use $|t_{j,\ell}|$; aggregated lag importances over ℓ are obtained by summation and normalization.

Relative-importance measures (random-forest–like scoring). We employ two complementary, scale-free scores:

$$s_j^{(p)} = \left[-\log_{10}(p_j)\right] \mathbf{1}\{p_j \leq \alpha\}, \ \text{Imp}_j^{(p)} = \frac{s_j^{(p)}}{\sum_{k \in S} s_k^{(p)}}, \tag{A.32}$$

$$s_j^{(\text{BIC})} = \max\{0, \Delta\text{BIC}_j\}, \ \text{Imp}_j^{(\text{BIC})} = \frac{s_j^{(\text{BIC})}}{\sum_k s_k^{(\text{BIC})}}, \tag{A.33}$$

where $S = \{j: p_j \leq \alpha\}$. Equation (A.32) mirrors random-forest practice by assigning positive mass only to statistically significant drivers and renormalizing their weights to sum to 1. Equation (A.33) interprets ΔBIC_j as an information-theoretic 'gain' and normalizes the positive gains across variables. In both cases, larger values indicate stronger evidence that x_j improves predictive content for y relative to other candidates.

Choice of a strict $\alpha = 0.01$ for short panels. In applied work with relatively short annual trade panels, the number of candidate variables and lags can be large relative to T. This inflates the number of simultaneous linear restrictions tested and the chance of spurious rejections. A stringent level $\alpha = 0.01$ controls Type I error more aggressively; under m independent tests, the family-wise error rate satisfies.

$$\text{FWER} \approx 1 - (1-\alpha)^m \leq m\alpha, \tag{A.34}$$

so reducing α curbs the expected number of false positives $m\alpha$ in small samples. This is particularly important when screening many x_j and lags ℓ.

Bayesian validation: Bayes factors and MCMC credible intervals. To corroborate classical rejections and guard against overfitting, we require Bayesian support in two forms.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

i. Bayes factor comparing full versus reduced models for x_j. Let $\mathcal{M}_{\text{full}}$ and $\mathcal{M}_{\text{red},j}$ be as above. The Bayes factor is

$$\mathrm{BF}_j = \frac{m(D|\mathcal{M}_{\text{full}})}{m(D|\mathcal{M}_{\text{red},j})} = \frac{\int L(\theta_{\text{full}};D)\,\pi(\theta_{\text{full}})\,d\theta_{\text{full}}}{\int L(\theta_{\text{red},j};D)\,\pi(\theta_{\text{red},j})\,d\theta_{\text{red},j}}, \tag{A.35}$$

with L the likelihood and $\pi(\cdot)$ prior densities. We adopt the widely used threshold $\mathrm{BF}_j \geq 5$ as 'substantial' evidence for the full model. In practice we also use the BIC approximation.

$$\log \mathrm{BF}_j \approx 1/2\left(\mathrm{BIC}(\mathcal{M}_{\text{red},j}) - \mathrm{BIC}(\mathcal{M}_{\text{full}})\right), \tag{A.36}$$

which coheres with $\Delta\mathrm{BIC}_j$ and is stable on short samples.

ii. Posterior exclusion of zero for at least one lag of x_j. Under (A.26) with a proper prior $\pi(\beta,\sigma^2)$, MCMC draws $\{\beta^{(r)},\sigma^{2(r)}\}_{r=1}^{R}$ target

$$p(\beta,\sigma^2|D) \propto L(\beta,\sigma^2;D)\,\pi(\beta,\sigma^2). \tag{A.37}$$

For each j we form credible intervals $\left[\beta_{j,\ell}^{L},\beta_{j,\ell}^{U}\right]$ at level $1-\gamma$ (e.g., 0.95). We declare Bayesian support for $x_j \to y$ if.

$$\exists\,\ell \in \{1,\ \ldots,p\} \text{ such that } 0 \notin \left[\beta_{j,\ell}^{L},\beta_{j,\ell}^{U}\right]. \tag{A.38}$$

Decision rule (intersection). A causal link is accepted only if follows:

$$\boxed{p_j \leq \alpha \wedge \mathrm{BF}_j \geq 5 \wedge (\text{A. 38}) \text{ holds}} \tag{A.39}$$

and then the expected delay $\widehat{\ell}_j$ is assigned by (A.31); otherwise $\widehat{\ell}_j := 0$.

Causal structure among the explanatory variables. Beyond $x_j \to y$, we characterize the internal dependence of the explanatory block X_t by constructing a directed graph $\mathcal{G} = (V,E,W)$ on $V = \{1,\ \ldots,d\}$:

$$(i \to k) \in E \Leftrightarrow x_i \text{ Granger} - \text{causes } x_k \text{ (same testing logic as above)}, \tag{A.40}$$

$$W_{ik} = \widehat{\ell}_{ik} \in \{1,\ \ldots,p\} \text{ (edge weight = estimated delay)}, \tag{A.41}$$

with $\widehat{\ell}_{ik}$ the lag maximizing $|\,t\,|$ in the x_k–equation when testing $x_i \to x_k$. This graph underpins the interpretation of y–drivers by revealing potential upstream/downstream channels among the predictors.

Short-sample regularization and implementation notes. In practice, to stabilize estimation with short T we (i) restrict p via data-driven selection or conservative caps; (ii) use information-criterion differentials $\Delta\mathrm{BIC}_j$ to grade evidence; and (iii) in the bivariate VAR workflow for each x_j use the Granger test p_j and the standardized t–statistics to form $\widehat{\ell}_j$. The relative-importance scores $\mathrm{Imp}_j^{(p)}$ and $\mathrm{Imp}_j^{(\mathrm{BIC})}$ are normalized to sum to 1 across significant variables, making them directly comparable across specifications and akin to random-forest variable-importance magnitudes.

Finally, we note two caveats. First, all procedures above detect only *linear* Granger-causal relations; nonlinear predictive effects are not identified by (A.26–A.27). Second, causal interpretation of $x_j \to y$ can be distorted when strong causal links exist among the predictors themselves (as reflected by $\mathcal{G}$). In subsequent sections we address these issues by employing random-forest models, which can flexibly capture nonlinearities and interactions and deliver alternative, stability-oriented importance measures that are less sensitive to linear-model misspecification.

Random Forest-Based Granger-Like Causality Analysis

Let $\{y_t\}_{t=1}^{n}$ denote the target (explained) time series and let $\{x_{j,t}\}_{t=1}^{n}$ for $j = 1,\ \ldots,p$ denote p candidate explanatory series. Fix a maximum lag $L \in \mathbb{N}$. For each time $t > L$, define the lagged covariate blocks.

$$u_t = (y_{t-1},y_{t-2},\ \ldots,y_{t-L}) \in \mathbb{R}^L, \qquad v_{j,t} = (x_{j,t-1},x_{j,t-2},\ \ldots,x_{j,t-L}) \in \mathbb{R}^L, \tag{A.42}$$

and the full predictor vector.

$$z_t = (u_t,v_{1,t},\ \ldots,v_{p,t}) \in \mathbb{R}^{L(p+1)}. \tag{A.43}$$

The effective sample is $\mathcal{T} = \{L+1,\ \ldots,n\}$ after trimming the first L indices.

We consider two random-forest regression models:

$$\text{Full model:}\quad y_t = f(z_t) + \varepsilon_t,$$
$$\text{Baseline (AR) model:}\quad y_t = g(u_t) + \eta_t,$$

where f and g are fitted by random forests on their respective features. A random-forest predictor with B trees has the form.

$$\widehat{f}(z) = \frac{1}{B}\sum_{b=1}^{B} T_b(z), \tag{A.44}$$

where T_b is the b-th regression tree trained on a bootstrap sample and using feature subsampling at splits. The AR baseline uses only u_t; if AR lags are disabled, the baseline reduces to a constant (the training mean).

To estimate out-of-sample (OOS) performance without temporal leakage, we partition $\mathcal{T}$ into K contiguous folds $F_1,\ \ldots,F_K$ of approximately equal size. For each $f = 2,\ \ldots,K$:

$$\text{Train on } \bigcup_{r=1}^{f-1} F_r, \qquad \text{test on } F_f. \tag{A.45}$$

Let $\widehat{g}^{(f)}$ and $\widehat{f}^{(f)}$ denote the baseline and full models trained for fold f. Define the OOS mean squared errors (MSEs)

$$\mathrm{MSE}_{\text{base}} = \frac{1}{N_{\text{test}}}\sum_{f=2}^{K}\sum_{t\in F_f}\left(y_t - \widehat{g}^{(f)}(u_t)\right)^2,$$
$$\mathrm{MSE}_{\text{full}} = \frac{1}{N_{\text{test}}}\sum_{f=2}^{K}\sum_{t\in F_f}\left(y_t - \widehat{f}^{(f)}(z_t)\right)^2, \tag{A.46}$$

where $N_{\text{test}} = \sum_{f=2}^{K} |\,F_f\,|$. Two summary measures reported are

$$\Delta_{\text{rel}} = 1 - \frac{\mathrm{MSE}_{\text{full}}}{\mathrm{MSE}_{\text{base}}} \quad (\text{relative MSE reduction; Granger} - \text{style incremental predictability}), \tag{A.47}$$

and the OOS R^2 against the test mean,

$$R^2_{\text{oos}} = 1 - \frac{\mathrm{MSE}_{\text{full}}}{\mathrm{Var}_{\text{test}}(y_t)}. \tag{A.48}$$

Let S index all features in the full model: the L lags of y and the L lags of each x_j (so $|\,S\,| = L(p+1)$). For a feature $s \in S$, let $z_t^{\pi(s)}$ denote z_t with the s-th coordinate permuted independently across the (out-of-bag or held-out) evaluation set.

The permutation (Breiman) importance for feature s is the OOB increase in prediction error when s is permuted:

$$\widehat{\mathrm{VI}}_{\text{perm}}(s) = \frac{1}{N_{\text{OOB}}}\sum_{t\in\text{OOB}}\left[\left(y_t - \widehat{f}\left(z_t^{\pi(s)}\right)\right)^2 - \left(y_t - \widehat{f}(z_t)\right)^2\right]. \tag{A.49}$$

In practice, the average is taken over OOB predictions aggregated across trees; positive values indicate that permuting s harms accuracy (thus s is important). Small negative values can occur due to sampling noise.

Alternatively, one may use the total decrease in node impurity (variance for regression) attributed to s and accumulated over all splits on s across all trees:

$$\widehat{\mathrm{VI}}_{\text{imp}}(s) = \frac{1}{B}\sum_{b=1}^{B}\sum_{\text{splits } v \text{ on } s}\left[\mathrm{Var}(\text{parent}_v) - \frac{n_L}{n_P}\mathrm{Var}(\text{left}_v) - \frac{n_R}{n_P}\mathrm{Var}(\text{right}_v)\right], \tag{A.50}$$

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

where (n_P, n_L, n_R) are sample sizes at the parent/left/right nodes.

In the implementation, importances are computed once by refitting the full model $\hat{f}$ on all available lagged observations $\{(z_t, y_t): t \in \mathcal{T}\}$ and extracting $\widehat{\mathrm{VI}}(s)$ for each feature s from the trained forest.

Index the lag-ℓ feature of variable x_j by (j, ℓ), $\ell = 1, \ldots, L$. Define:

$$\widehat{\mathrm{VI}}(j,\ell) = \widehat{\mathrm{VI}}(s=(j,\ell)), \widehat{\mathrm{VI}}_j = \sum_{\ell=1}^{L} \widehat{\mathrm{VI}}(j,\ell) \ \text{(variable-level importance)}. \tag{A.51}$$

A relative variable importance (over X only) is obtained via normalization.

$$\widetilde{\mathrm{VI}}_j = \frac{\widehat{\mathrm{VI}}_j}{\sum_{m=1}^{p} \widehat{\mathrm{VI}}_m}, \tag{A.52}$$

whenever the denominator is positive (otherwise normalization is left undefined). Analogously, the global importance of lag ℓ across all X variables is.

$$\widehat{\mathrm{VI}}^{\mathrm{lag}}(\ell) = \sum_{j=1}^{p} \widehat{\mathrm{VI}}(j,\ell), \qquad \widetilde{\mathrm{VI}}^{\mathrm{lag}}(\ell) = \frac{\widehat{\mathrm{VI}}^{\mathrm{lag}}(\ell)}{\sum_{r=1}^{L} \widehat{\mathrm{VI}}^{\mathrm{lag}}(r)}. \tag{A.53}$$

Importances for the autoregressive lags of y are computed in the same way using the features $\{y_{t-\ell}\}_{\ell=1}^{L}$.

For each explanatory variable x_j, the most likely (dominant) lag is chosen as the lag with the largest importance:

$$\hat{\ell}_j \in \arg\max_{\ell \in \{1,\ldots,L\}} \widehat{\mathrm{VI}}(j,\ell), \tag{A.54}$$

with ties broken in favour of the smallest lag. A variable-level drop rule declares 'no Granger-style predictability' if the largest lag-importance fails to exceed a nonnegative threshold τ_0:

$$\max_{\ell} \widehat{\mathrm{VI}}(j,\ell) \le \tau_0 \ \Rightarrow \ \text{drop } x_j \text{ and set expected lag } \hat{\tau}_j = 0, \tag{A.55}$$

otherwise $\hat{\tau}_j = \hat{\ell}_j$. In the implementation, $\tau_0 = 0$ by default (appropriate for permutation importance where negative values can occur). For additional interpretability, a within-variable 'lag share' is also reported:

$$\mathrm{lag_share}_j = \frac{\widehat{\mathrm{VI}}\left(j, \hat{\ell}_j\right)}{\sum_{\ell=1}^{L} \widehat{\mathrm{VI}}(j,\ell)} \quad \text{(defined when the denominator is positive)}. \tag{A.56}$$

The procedure operationalizes Granger causality in a predictive sense: x_j is said to Granger-cause y (at some lag) if including $\{x_{j,t-\ell}\}_{\ell=1}^{L}$ alongside y's own lags improves the OOS prediction of y_t, as evidenced by (i) a positive and dominant $\widehat{\mathrm{VI}}(j,\ell)$ for some ℓ, and/or (ii) a reduction in OOS error relative to the AR baseline ($\Delta_{\mathrm{rel}} > 0$). This is not a statement of structural (interventional) causality; it detects incremental predictability conditional on past information.

The steps of Random-Based Causality Analysis

1. Build the lagged dataset $\{(z_t, y_t)\}_{t\in\mathcal{T}}$ for a chosen L.
2. Blocked CV: for $f = 2, \ldots, K$, train $\hat{g}^{(f)}$ on u_t and $\hat{f}^{(f)}$ on z_t using $\bigcup_{r<f} F_r$; score OOS on F_f to compute $\mathrm{MSE}_{\mathrm{base}}$, $\mathrm{MSE}_{\mathrm{full}}$, Δ_{rel}, and R^2_{oos}.
3. Fit the final full forest $\hat{f}$ on all $\{(z_t, y_t)\}$ and compute feature importances $\widehat{\mathrm{VI}}(s)$ (permutation or impurity).
4. Aggregate to $\widehat{\mathrm{VI}}_j$, $\widetilde{\mathrm{VI}}_j$, $\widehat{\mathrm{VI}}(j,\ell)$, and $\widehat{\mathrm{VI}}^{\mathrm{lag}}(\ell)$; select expected lags $\hat{\tau}_j$ with the drop rule based on τ_0.

It should be noted that, permutation importances can be slightly negative; the default drop threshold $\tau_0 = 0$ treats such cases as non-evidence for predictability. Group-level normalizations (e.g., relative variable importances across X) are performed only when the relevant totals are positive. Blocked CV (growing training window) avoids look-ahead bias and mimics realistic forecasting evaluation.

A variable x_j is declared to exhibit RF-based Granger causality to y if and only if:

$$\boxed{\widehat{\mathrm{VI}}_j > 0 \wedge \max_{\ell} \widehat{\mathrm{VI}}(j,\ell) > \max_{k,\ell} \widehat{\mathrm{VI}}\left(\tilde{x}_{j,\ell}^{(k)}\right)} \tag{A.57}$$

where the first condition ensures positive aggregate importance and the second requires the variable's best lag to exceed all shadow feature importances. The expected lag is then set to $\hat{\tau}_j = \arg\max_{\ell} \widehat{\mathrm{VI}}(j,\ell)$.

Panel Version of Granger-Causality Tests

Let $Y_{j,t}$ denote the dependent variable for sector j in period t, where $j \in \{1, \ldots, J\}$ and $t \in \{1, \ldots, T\}$. Let $X_{i,j,t}$ represent explanatory variable i for sector j in period t, where $i \in \{1, \ldots, M\}$. The panel Granger causality framework tests whether past values of $X_{i,j,t}$ contain information that helps predict $Y_{j,t}$ beyond what is already captured by past values of $Y_{j,t}$ itself.

Traditional Panel Granger Causality

Following the seminal work of Granger (1969), Granger causality in a panel context (Holtz-Eakin et al. 1988; Dumitrescu and Hurlin 2012) is assessed by comparing a *restricted* autoregressive model against an *unrestricted* model that includes lagged values of the candidate causal variable.

For each sector j and explanatory variable i, the restricted model is:

$$Y_{j,t} = \alpha_j + \sum_{\ell=1}^{p} \beta_{\ell,j} Y_{j,t-\ell} + \varepsilon_{j,t}, \tag{A.58}$$

where p is the lag order, α_j is a sector-specific intercept, and $\varepsilon_{j,t}$ is an error term. The unrestricted model augments (A.58) with lagged values of $X_{i,j,t}$:

$$Y_{j,t} = \alpha_j + \sum_{\ell=1}^{p} \beta_{\ell,j} Y_{j,t-\ell} + \sum_{\ell=1}^{p} \delta_{\ell,i,j} X_{i,j,t-\ell} + u_{j,t}. \tag{A.59}$$

The null hypothesis $H_0: \delta_{1,i,j} = \delta_{2,i,j} = \cdots = \delta_{p,i,j} = 0$ states that $X_{i,j,t}$ does *not* Granger-cause $Y_{j,t}$. This hypothesis is tested using a Wald test (F-test) that compares the residual sum of squares from (A.58) and (A.59):

$$F_{i,j,p} = \frac{\left(\mathrm{RSS}_{\mathrm{restricted}} - \mathrm{RSS}_{\mathrm{unrestricted}}\right)/p}{\mathrm{RSS}_{\mathrm{unrestricted}}/(T-2p-1)}, \tag{A.60}$$

where RSS denotes the residual sum of squares. Under the null and standard regularity conditions, $F_{i,j,p} \sim F(p, T-2p-1)$. Rejection at significance level α (e.g., $\alpha = 0.01$) provides evidence that $X_{i,j,t}$ Granger-causes $Y_{j,t}$ at lag p.

For pooled (panel-wide) inference, we stack observations across sectors and estimate a panel model with sector fixed effects:

$$Y_{j,t} = \alpha_j + \sum_{\ell=1}^{p} \beta_{\ell} Y_{j,t-\ell} + \sum_{\ell=1}^{p} \delta_{\ell,i} X_{i,j,t-\ell} + u_{j,t}, \tag{A.61}$$

where the fixed effects α_j control for unobserved sector-specific heterogeneity. The panel Wald test then evaluates whether $\delta_{1,i} = \cdots = \delta_{p,i} = 0$ across all sectors jointly (Dumitrescu and Hurlin 2012).

Lag selection. We test lag orders $p \in \{1, \ldots, p_{\max}\}$ sequentially and report the *minimum* lag $\hat{p}_{i,j}$ at which the null is rejected. If no lag achieves significance, we set $\hat{p}_{i,j} = 0$, indicating no detectable Granger causality. The resulting matrix $\mathbf{P} = [\hat{p}_{i,j}]_{M\times(J+1)}$ summarizes the causal structure: columns $1, \ldots, J$ correspond to individual sectors, and column $J+1$ to the pooled panel.

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Bayesian Panel Granger Causality

The Bayesian approach replaces hypothesis testing with model comparison via Bayes factors (Koop and Korobilis 2010, 2013). For each lag p, we compute the marginal likelihoods of the restricted model $\mathcal{M}_p^{(R)}$ (Equation A.58) and the unrestricted model $\mathcal{M}_p^{(U)}$ (Equation A.59) using Jeffreys–Zellner–Siow (JZS) priors on the regression coefficients (Liang et al. 2008; Rouder et al. 2012). The Bayes factor in favour of the unrestricted model is:

$$\mathrm{BF}_{i,j,p} = \frac{p\left(\mathbf{Y}_j \mid \mathcal{M}_p^{(U)}\right)}{p\left(\mathbf{Y}_j \mid \mathcal{M}_p^{(R)}\right)}, \tag{A.62}$$

where $\mathbf{Y}_j = \left(Y_{j,p+1}, \ldots, Y_{j,T}\right)^\top$ is the observed data vector for sector j. Larger values of $\mathrm{BF}_{i,j,p}$ indicate stronger evidence that including lagged $X_{i,j,t}$ improves predictions of $Y_{j,t}$.

Following established conventions (Kass and Raftery 1995; Jeffreys 1961), we interpret the strength of evidence as follows:

$\mathrm{BF}_{i,j,p} \geq 3$ positive evidence for Granger causality,

$\mathrm{BF}_{i,j,p} \geq 5$ strong evidence (default threshold),

$\mathrm{BF}_{i,j,p} \geq 10$ very strong evidence,

$\mathrm{BF}_{i,j,p} \geq 30$ decisive evidence.

For each explanatory variable i and sector j, we compute $\mathrm{BF}_{i,j,p}$ for all $p \in \{1, \ldots, p_{\max}\}$ and report the minimum lag $\hat{p}_{i,j}$ satisfying $\mathrm{BF}_{i,j,p} \geq \mathrm{BF}_{\min}$ (with $\mathrm{BF}_{\min} = 5$ by default). If no lag meets this criterion, we set $\hat{p}_{i,j} = 0$.

For the pooled panel, we extend (A.61) by treating sector fixed effects α_j as random effects and computing Bayes factors that integrate over the posterior distribution of $(\alpha_1, \ldots, \alpha_J)$. This accounts for cross-sectional heterogeneity while assessing Granger causality at the panel level.

The Bayesian framework avoids multiple-testing issues inherent in sequential hypothesis testing and provides a continuous measure of evidence strength. It naturally incorporates prior information and yields probabilistic statements about model plausibility rather than binary reject/fail-to-reject decisions.

Random Forest Panel Granger Causality

ML methods can capture nonlinear Granger causality that linear models may miss. We employ Random Forests (Breiman 2001), a nonparametric ensemble method, combined with a Boruta-style variable selection procedure (Kursa and Rudnicki 2010) to identify relevant lagged predictors.

For lag p, we construct a training dataset:

$$\mathcal{D}_{j,p} = \left\{\left(Y_{j,t-1}, \ldots, Y_{j,t-p}, X_{i,j,t-1}, \ldots, X_{i,j,t-p}; Y_{j,t}\right)\right\}_{t=p+1}^{T}. \tag{A.63}$$

A Random Forest $\mathcal{F}_{j,p}$ is trained to predict $Y_{j,t}$ from the lagged features $\left\{Y_{j,t-\ell}, X_{i,j,t-\ell}\right\}_{\ell=1}^{p}$ via an ensemble of decision trees. Each tree is grown on a bootstrap sample, and at each node, a random subset of features is considered for splitting. The final prediction is the average (for regression) over all trees (Breiman 2001).

Variable importance and shadow features To assess whether the lagged $X_{i,j,t-\ell}$ variables contribute meaningful predictive power, we implement a Boruta-inspired shadow feature approach (Kursa and Rudnicki 2010):

1. For each feature $Z \in \left\{Y_{j,t-1}, \ldots, Y_{j,t-p}, X_{i,j,t-1}, \ldots, X_{i,j,t-p}\right\}$, create *K shadow features* $\tilde{Z}^{(1)}, \ldots, \tilde{Z}^{(K)}$ by randomly permuting the values of Z across observations. Shadow features have no true predictive relationship with $Y_{j,t}$ by construction.
2. Train the Random Forest on the augmented dataset that includes both original and shadow features.
3. Compute permutation-based variable importance [@breiman2001] for each feature. Permutation importance measures the increase in prediction error when the feature's values are randomly shuffled, quantifying its contribution to model accuracy.
4. For explanatory variable i, compute the average importance of its lagged features:

$$I_{i,j,p} = \frac{1}{p} \sum_{\ell=1}^{p} \mathrm{Imp}\left(X_{i,j,t-\ell}\right), \tag{A.64}$$

where $\mathrm{Imp}\left(X_{i,j,t-\ell}\right)$ denotes the permutation importance of $X_{i,j,t-\ell}$.

5. Compute the maximum shadow importance:

$$I_{\max j,p}^{\text{shadow}} = \max_{\ell,k} \mathrm{Imp}\left(\tilde{X}_{i,j,t-\ell}^{(k)}\right). \tag{A.65}$$

Variable $X_{i,j,t}$ is deemed to Granger-cause $Y_{j,t}$ at lag p if and only if $I_{i,j,p} > I_{\max j,p}^{\text{shadow}}$—that is, its importance exceeds that of all shadow (purely random) features.

Lag selection. We compute $I_{i,j,p}$ and $I_{\max j,p}^{\text{shadow}}$ for all $p \in \{1, \ldots, p_{\max}\}$ and identify the lag $\hat{p}_{i,j}$ at which $I_{i,j,p}$ is maximized *and* satisfies $I_{i,j,\hat{p}_{i,j}} > I_{\max j,\hat{p}_{i,j}}^{\text{shadow}}$.

If no lag meets this condition, we set $\hat{p}_{i,j} = 0$, indicating that $X_{i,j,t}$ does not provide predictive power beyond random chance.

For the pooled panel, we stack observations across sectors and include sector indicators as categorical features in the Random Forest. The same shadow-feature procedure is applied to the pooled data, yielding a panel-level lag estimate $\hat{p}_{i,J+1}$.

The Random Forest approach accommodates nonlinear and interaction effects without prespecification, is robust to outliers and distributional violations, and naturally handles high-dimensional predictor spaces. The shadow-feature method provides a principled threshold for distinguishing signal from noise, akin to the Boruta algorithm's strategy for all-relevant feature selection (Kursa and Rudnicki 2010).

14679701, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/twec.70148 by Cochrane Hungary, Wiley Online Library on [10/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License