



Universiteit  
Leiden  
The Netherlands

## Functional structure in production networks

Mattsson, E.S.C.; Takes, F.W.; Heemskerk, E.M.; Diks, C.; Buiten, G.; Faber, A.; Sloot, P.M.A.

### Citation

Mattsson, E. S. C., Takes, F. W., Heemskerk, E. M., Diks, C., Buiten, G., Faber, A., & Sloot, P. M. A. (2021). Functional structure in production networks. *Frontiers In Big Data*, 4, 1-13. doi:10.3389/fdata.2021.666712

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3276998>

**Note:** To cite this publication please use the final published version (if applicable).



# Functional Structure in Production Networks

Carolina E. S. Mattsson<sup>1,2\*</sup>, Frank W. Takes<sup>1,3</sup>, Eelke M. Heemskerk<sup>3,4</sup>, Cees Diks<sup>5,6</sup>, Gert Buiten<sup>7</sup>, Albert Faber<sup>8</sup> and Peter M. A. Sloot<sup>9,10,11,12,13</sup>

<sup>1</sup> Computational Network Science Lab, Leiden Institute of Advanced Computer Science, Leiden University, Leiden, Netherlands, <sup>2</sup> Network Science Institute, Boston, MA, United States, <sup>3</sup> CORPNET, University of Amsterdam, Amsterdam, Netherlands, <sup>4</sup> Department of Political Science, University of Amsterdam, Amsterdam, Netherlands, <sup>5</sup> Faculty Economics and Business, University of Amsterdam, Amsterdam, Netherlands, <sup>6</sup> Tinbergen Institute, Amsterdam, Netherlands, <sup>7</sup> Statistics Netherlands, The Hague, Netherlands, <sup>8</sup> Ministry of Economic Affairs & Climate, The Hague, Netherlands, <sup>9</sup> Computational Science Lab, Faculty of Science, University of Amsterdam, Amsterdam, Netherlands, <sup>10</sup> Institute for Advanced Study, University of Amsterdam, Amsterdam, Netherlands, <sup>11</sup> Complexity Institute, Nanyang Technological University, Singapore, Singapore, <sup>12</sup> Complexity Science Hub Vienna, Vienna, Austria, <sup>13</sup> National Center for Cognitive Research, ITMO University, Saint Petersburg, Russia

## OPEN ACCESS

### Edited by:

Michele Coscia,  
IT University of Copenhagen, Denmark

### Reviewed by:

Zafer Kanik,  
University of Glasgow,  
United Kingdom  
Benjamin Renoust,  
Osaka University, Japan

### \*Correspondence:

Carolina E. S. Mattsson  
e.s.c.mattsson@liacs.leidenuniv.nl;  
mattsson.c@northeastern.edu

### Specialty section:

This article was submitted to  
Big Data Networks,  
a section of the journal  
Frontiers in Big Data

**Received:** 10 February 2021

**Accepted:** 19 April 2021

**Published:** 21 May 2021

### Citation:

Mattsson CES, Takes FW, Heemskerk EM, Diks C, Buiten G, Faber A and Sloot PMA (2021) Functional Structure in Production Networks. *Front. Big Data* 4:666712. doi: 10.3389/fdata.2021.666712

Production networks are integral to economic dynamics, yet dis-aggregated network data on inter-firm trade is rarely collected and often proprietary. Here we situate company-level production networks within a wider space of networks that are different in nature, but similar in local connectivity structure. Through this lens, we study a regional and a national network of inferred trade relationships reconstructed from Dutch national economic statistics and re-interpret prior empirical findings. We find that company-level production networks have so-called functional structure, as previously identified in protein-protein interaction (PPI) networks. Functional networks are distinctive in their over-representation of closed squares, which we quantify using an existing measure called spectral bipartivity. Shared local connectivity structure lets us ferry insights between domains. PPI networks are shaped by complementarity, rather than homophily, and we use multi-layer directed configuration models to show that this principle explains the emergence of functional structure in production networks. Companies are especially similar to their close competitors, not to their trading partners. Our findings have practical implications for the analysis of production networks and give us precise terms for the local structural features that may be key to understanding their routine function, failure, and growth.

**Keywords:** production networks, inter-firm networks, complexity economics, economic statistics, trade linkages, functional networks, bipartivity

## 1. INTRODUCTION

It has become established knowledge within complexity economics (Arthur, 2021) that network structure affects economic dynamics over the short-, medium-, and long-term. Cascading processes over networks have been used to model supply disruptions that propagate in a matter of days or weeks (Burkholz, 2016; Inoue and Todo, 2019). Trade linkages have been used to explain aggregate fluctuations in business activity that play out over months or years (Acemoglu et al., 2012; Carvalho and Tahbaz-Salehi, 2018). Structural changes happen over decades and we know that national

growth trajectories are strongly affected by the network structure of economic activity (Hausmann and Hidalgo, 2013; McNerney et al., 2018). The routine function, failure, and growth of economic production networks are at the heart of these dynamics.

The mechanisms underlying dynamics on and of production networks are thought to operate at the level of trade relationships among individual companies (see Hazama and Uesugi, 2017; Carvalho and Tahbaz-Salehi, 2018; Inoue and Todo, 2019). However, our understanding of company-level production networks is limited; empirical network data on customer-supplier ties is not often available and, when it is, generally proprietary (Fujiwara and Aoyama, 2010; Ohnishi et al., 2010; Magerman et al., 2015). Moreover, findings based on such data remain difficult to interpret in the context of wider research on production networks because they combine local and national scales. Trade relationships are more often considered *either* in detailed, local case studies or as aggregated trade linkages among sectors, industries, and countries based in officially-prepared macroeconomic statistics (Uzzi, 1997; Coenen et al., 2010; Acemoglu et al., 2012; Miller and Temurshoev, 2017; McNerney et al., 2018).

In this work, we situate company-level production networks within a wider space of networks that are different in nature, but similar in structure. For this we develop a typology based on well-known and recognizable *local connectivity structures*. Random networks, social networks, and two-mode networks are the most important and well-known connectivity types and they form the basis for our typology (Borgatti and Everett, 1997; Newman et al., 2001; Rivera et al., 2010). Commonalities among networks with the same local connectivity type present opportunities to use established knowledge from one domain to better understand networks in another. Specifically, a new network type with a distinctive local connectivity structure has recently been identified in work on protein-protein interaction (PPI) networks (Kovács et al., 2019; Kitsak, 2020). This adds so-called functional networks to our typology. We highlight several existing empirical findings (Fujiwara and Aoyama, 2010; Ohnishi et al., 2010) to suggest that company-level production networks are best characterized as functional networks and go on to explore this hypothesis.

Specifically, we consider a regional and a national company-level production network reconstructed from Dutch national economic statistics. As detailed in Hooijmaaijers and Buiten (2019), Statistics Netherlands (CBS) has used official statistics to systematically infer customer-supplier ties among companies in the Netherlands for each of 677 product groups (e.g., “Electricity,” “Fertilizer,” “Shipping services,” etc.). Starting from this multi-layer network, we analyze the local connectivity structure of the network of unique (inferred) trade relationships among companies with 5+ employees within Zeeland province and the whole of the Netherlands. We then generalize the CBS reconstruction process to produce ensembles of networks of (hypothetical) trade relationships among the Zeeland companies with 5+ employees; this lets us study how local connectivity structure emerges in company-level production networks.

To assess whether networks have functional structure, we re-purpose an existing measure that captures a distinctive feature of local connectivity in functional networks. Link prediction performance on PPI networks suggests that three-step closure is more common than two-step closure (Kovács et al., 2019). That is, closed squares are more prominent than are closed triangles. We measure this using spectral bipartivity, which quantifies the abundance of even vs. odd cycles in a network’s local connectivity structure (Estrada and Gómez-Gardeñes, 2016), in comparison to random expectation; functional networks have higher-than-random spectral bipartivity.

We do indeed find functional structure in both the national and the regional company-level production networks. Compared to randomized versions of itself, the reconstructed regional network has significantly higher spectral bipartivity [Kolmogorov-Smirnov (KS) statistic 1.0,  $N_1 = 1$ ,  $N_2 = 1,000$ ,  $p < 0.001$ ]. The value itself remains much smaller than 1, indicating that the network structure is definitely not bipartite as would be the case if it were a two-mode network. Small values then let us approximate the logit-transformed spectral bipartivity and extend our results to the larger Netherlands network. This measure is again higher than expected ( $KS = 1.0$ ,  $N_1 = 1$ ,  $N_2 = 25$ ,  $p < 0.04$ ) indicating functional structure.

In the literature on protein-protein interactions, network structure is thought to be shaped by the principle of *complementarity* as opposed to the principle of *homophily* well-known to shape social networks (McPherson et al., 2001; Kitsak, 2020). Node complementarity in PPI networks reflects the practical fact that proteins physically bind with one another at compatible binding sites (Kovács et al., 2019). This is a useful concept for company-level production networks, as well, since trade relationships imply the exchange of some product between companies that hold complementary roles as customer and supplier. Trade compatibility imposes a specific constraint on the formation of trade relationships that might explain the emergence of functional structure in production networks.

To test this explanation, we use multi-layer directed configuration models to generate ensembles of networks made up of trade-compatible relationships within Zeeland. In each layer, we make a random selection of the possible ties between customers and suppliers of products in a product group. By defining trade compatibility according to progressively more detailed product categorizations, we ramp up the strictness of the constraint. These generated networks are then compared against their randomized version where customer-supplier complementarity has been broken. Our analysis finds that imposing more stringent complementarity introduces consistently and significantly more functional structure ( $KS = 1.0$ ,  $N_1 = 25$ ,  $N_2 = 25$ ,  $p < 10^{-14}$ ).

Identifying functional structure in company-level production networks has practical implications for the analysis of such networks and wider implications for network science and complexity economics. Under the logic of functional networks, companies trade with complementary others and are especially similar to their close competitors, not their trading partners. This also implies that production networks are markedly different from social and economic networks driven by

homophily; they are more comparable to PPI networks and food webs. More generally, we have demonstrated the usefulness of network categorization according to a typology of local connectivity structure. In our case, identifying functional structure lets us bring in precise language from previously unrelated domains to describe defining features of company-level production networks: “node complementarity,” “closed squares,” and “functional modules” (Barabási and Oltvai, 2004; Kovács et al., 2019; Kitsak, 2020). Studying these structural features of production networks could be key to furthering our understanding of their routine function, failure, and growth.

The remainder of the paper is structured as follows. In section 2, we describe several types of networks with distinct local connectivity structure and discuss in detail highly relevant prior work. Section 3 describes the company-level production networks that we study and the methods we use to characterize their local connectivity structure. In section 4 we present our findings and in section 5 we discuss their implications.

## 2. THEORY

Networks studied in different domains can nonetheless be similar in structure. Section 2.1 develops a network typology based around the local connectivity structures of random, social, and two-mode networks. To this typology we add recently identified so-called “functional” networks. In section 2.2 we advance the hypothesis that company-level production networks are best characterized as functional networks.

### 2.1. Network Typology

Different types of networks can have systematically different structural properties. Many network analysis tools have been developed especially for use with networks of some particular type (Opsahl, 2013; Masuda et al., 2018). The same algorithms can have markedly different performance across network types (Ghasemian et al., 2020) and the same measures will often have different typical ranges (Newman, 2003b; Costa et al., 2007). Network measures also tend to be correlated (Jamakovic and Uhlig, 2008; Bounova and de Weck, 2012) with local network features affecting global ones (Colomer-de Simón et al., 2013; Jamakovic et al., 2015; Asikainen et al., 2020). As such, we use the following network typology based in recognizably distinct *local connectivity structures*.

#### 2.1.1. Random Networks

Random networks are those where the set of existing links might have come to occur by chance. Several celebrated network properties, such as the emergence of a *giant component* and the *small world* property, are identifiable already in random networks (Bollobás, 2001; Newman et al., 2001). Other common network properties, such as the existence of *hubs* or *communities*, can be introduced using block-wise random networks (Holland et al., 1983; Newman et al., 2001; Karrer and Newman, 2011). This flexibility makes random networks especially useful as a baseline comparison for empirical network data and generative network models (Costa et al., 2007). Some measures, such as

modularity (Newman, 2006) and degree assortativity (Pastor-Satorras et al., 2001; Newman, 2003a), directly incorporate a comparison to random expectation.

#### 2.1.2. Two-Mode Networks

Two-mode networks are those where the links are affiliations between categorically different nodes. For instance, directors are affiliated with the corporate boards on which they serve (Mizruchi, 1996; Seierstad and Opsahl, 2011; Takes and Heemskerck, 2016; Valeeva et al., 2020). Two-mode networks are *bipartite* in that the nodes can be separated into two groups where links exist between, but not within, groups (Borgatti and Everett, 1997; Holme et al., 2003). This is a hard constraint on the existence of network links that affects the interpretation of any network-structural property. As such, there are specific network analysis techniques and particular versions of centrality, modularity, closure, and other measures developed for two-mode networks (Borgatti and Everett, 1997; Barber, 2007; Zhou et al., 2007; Opsahl, 2013; Berardo, 2014; Kunegis, 2015).

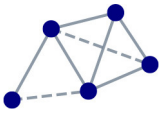
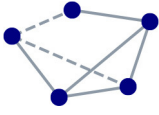
#### 2.1.3. Social Networks

Social networks are those where links exist between nodes who associate with one another in some sense. This is a soft constraint on the formation of links that affects many network properties. Social networks have high *clustering* as various social dynamics are known to generate *triadic closure* (see Rivera et al., 2010). They are typically shaped by *homophily* leading to high *assortativity*, where nodes who associate with one another tend to be similar in some sense (McPherson et al., 2001; Newman, 2003a). Social networks frequently contain interpretable *communities* (Blondel et al., 2008; Davis et al., 2009) and are often assortative in degree (Johnson et al., 2010). It has been suggested that the tendency of nodes to form communities may be a derivative feature of triadic closure operating locally (Colomer-de Simón et al., 2013; Jamakovic et al., 2015). The over-representation of triangles also allows link prediction algorithms that operate on a two-step basis (L2) to perform well on social networks (Ghasemian et al., 2020). Degree assortativity may also be related to triadic closure, homophily, and community structure (Newman and Park, 2003; Asikainen et al., 2020).

#### 2.1.4. Functional Networks

Functional networks are those where links form between nodes that complement one another in fulfilling some function. This network type has been identified in work on protein-protein interaction (PPI) networks; proteins bind not with similar others but with those that have a binding site physically compatible with their own (Kovács et al., 2019). These networks are shaped by the principle of *complementarity*, a different soft constraint that likely also affects many, related, network properties (Kitsak, 2020). Kovács et al. (2019) establishes that L2 heuristics under-perform in link prediction on PPI networks while those operating on a three-step basis (L3) are more accurate. In analogy with social networks, this might suggest that functional dynamics generate *tetradic closure*. Meso-scale structures in PPI networks have been termed *functional modules* as they are often interpretable as key

**TABLE 1** | Key differences between social and functional networks.

	Toy network	Example	Principle	Closure	Meso-scale
Social		People are often friends with others similar to themselves, and a pair of friends likely also have friends in common.	Homophily	Triadic	Community
Functional		Proteins interact at a physical binding site; proteins are likely to bind if one is similar to the other's other partners.	Complementarity	Tetradic	Module

to some higher-level cellular function (Barabási and Oltvai, 2004; Chen and Yuan, 2006; Ghiassian et al., 2015). *Disassortativity* in degree may be an expected property of functional networks more broadly (Johnson et al., 2010; Barabási and Pósfai, 2016). Food webs and interdisciplinary collaboration networks have been proposed as additional domains where the same structural patterns are likely to be found (Kitsak, 2020).

**Table 1** highlights the difference in local connectivity structure between social and functional network types, with links likely to form shown using dashed lines. In social networks links form between nodes who associate with one another, nodes who associate with one another are similar in some sense, and various social dynamics generate closed triangles. Social networks thus have a higher clustering coefficient than random networks. The analogous intuition for functional networks is that links form between nodes in fulfilling their function, that nodes who complement the same partners are similar in some sense, and that various functional dynamics generate closed squares. Functional networks would score higher than random networks on measures that capture this aspect of local connectivity structure.

## 2.2. Features of Company-Level Production Networks

Characterizing local connectivity structure directly, as we do here, offers a new lens through which to consider known structural features of networks. Findings within existing studies of empirical company-level production networks are thus exceedingly relevant. Here we discuss in detail two papers that describe the structure of a Japanese inter-firm network produced by Tokyo Shoko Research Ltd (Fujiwara and Aoyama, 2010; Ohnishi et al., 2010). This is a company-level production network where nodes are Japanese firms and links correspond to trade relationships in materials or services; financial relationships are downplayed. Several findings in these papers let us form the hypothesis that company-level production networks have functional structure.

Ohnishi et al. (2010) perform a comprehensive analysis of three-node motifs on the Japanese inter-firm network from 2005. On a simple, directed network there are thirteen possible three-node motifs and this paper presents their empirical prevalence compared to random expectation (Ohnishi et al., 2010, Figures 4, 5). Recall that closed squares, but not closed triangles, feature prominently in the local connectivity structure of functional

networks. In this paper, the most substantially over-represented motifs in the Japanese inter-firm network are two-link, V-shaped motifs; these are the most compatible with closed squares. Three-link loops forming closed triangles are found to be the most substantially under-represented. These findings can be interpreted as empirical evidence of functional structure in this company-level production network.

Fujiwara and Aoyama (2010) conduct a multi-pronged network analysis of the Japanese inter-firm network compiled in September 2006. Three of their findings are especially relevant with respect to our typology of local connectivity structure. First, this paper shows the network exhibits disassortativity in degree (Fujiwara and Aoyama, 2010, Figure 3). Recall that disassortativity may be an expected feature of functional networks as is assortativity for social networks. Second, this paper performs a detailed analysis of meso-scale structure within the sub-network defined by firms in the manufacturing sector. Using modularity maximization they identify a large number of small groups with a striking qualitative interpretation: “From the database of the information on the firms, we found that many of those small communities are each located in same geographical areas forming specialized production flows. An example is a small group of flour-maker, noodle-foods producers, bakeries, and packing/labeling companies in a rural area” (Fujiwara and Aoyama, 2010, p. 570). While these groups are referred to as “communities,” they might be better understood as so-called “functional modules” in that they are key to the functioning of economic production at a higher level. Finally, this paper finds locally bipartite structure within prominent (sub-)industries also identified by modularity maximization. In each (sub-)industry there is a handful of large, recognizable firms who are not often directly linked, as they are competitors, but have many suppliers (and customers) in common (Fujiwara and Aoyama, 2010, p. 570). Locally bipartite meso-scale structure means there are many closed squares, enough that these groups are picked up by modularity maximization. Closed squares, functional modules, and disassortativity are all hallmarks of functional networks.

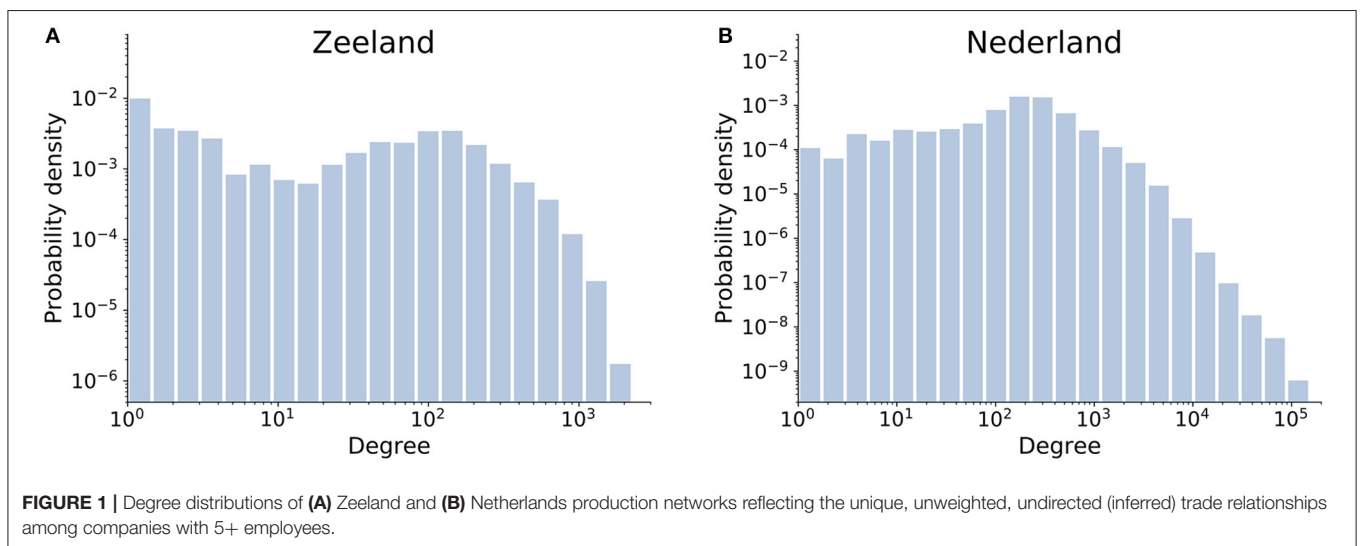
## 3. DATA AND METHODS

This paper makes use of a network dataset produced by Statistics Netherlands (CBS) to explore our hypothesis that company-level production networks have functional structure. In section 3.1 we

**TABLE 2** | Description of the data provided by CBS.

	Nodes	Layers	Multi-edges	Simple edges	GCC Nodes	GCC Multi-edges	GCC Simple edges
<b>NETHERLANDS</b>							
All	875,222	677	310,324,477	195,903,806	875,222	310,324,477	195,903,806
5+	102,461	677	116,652,466	50,930,077	102,461	116,652,466	50,930,077
<b>ZEELAND</b>							
All	18,398	667	3,500,797	2,143,412	18,337	3,500,762	2,143,379
5+	2,497	652	848,015	334,334	2,497	848,015	334,334

Highlighted are the reconstructed networks for the Netherlands and Zeeland province above our company size threshold (measured using the number of employees). Also shown are the statistics for the giant connected component (GCC) of these networks.



describe the data itself and how we make use of it. Section 3.2 presents our methods, describing how we characterize local connectivity structure.

### 3.1. Data

CBS has produced a network dataset of systematically inferred customer-supplier ties among Dutch companies (Hooijmaaijers and Buiten, 2019). In section 3.1.1 we describe how we define a national and a regional company-level production network from this data; we also briefly relay how this data was produced. Section 3.1.2 details how we generalize the CBS reconstruction process, in this work, to construct ensembles of many possible customer-supplier networks within Zeeland province.

#### 3.1.1. National and Regional Company-Level Production Networks

We begin with an existing dataset of inferred customer-supplier ties produced by researchers at CBS (Hooijmaaijers and Buiten, 2019). This network has 677 layers, each corresponding to domestic trade in a particular product group. Product groups include “Electricity,” “Fertilizer,” “Shipping services,” and “Accounting & tax administration” as in the Dutch implementation of the European Classification of Products by Activity (Eurostat, 2008b, CPA 2008). Companies likely to supply

products from a product group were matched with companies likely to use such products, with inferences based in Dutch economic statistics for the year 2012. Customer-supplier ties in each layer should be understood to reflect a selection of trading pairs among sets of likely customers and suppliers for products in that product group. **Table 2** describes the reconstructed network of inferred customer-supplier ties within the Netherlands and within Zeeland province; self-loops have been removed. Zeeland is a low-lying delta region home to 381,407 people in 2012 (Statline, 2012).

Based on the CBS data, this work defines a national and a regional company-level production network. For this, we first consider the sub-network among companies with five or more employees within Zeeland province and the whole of the Netherlands. We then focus on the “simple” versions of these multi-layer, directed, customer-supplier networks. Simple networks are those with unique, unweighted, undirected links. That is, we place an (inferred) trade relationship between any two companies with an inferred customer-supplier tie, in either direction, in any of the product groups. To give an idea of the heterogeneity in our version of these networks, **Figure 1** describes their degree distributions on a log-log scale. Filtering out companies with few employees is done as the underlying, company-level statistics are deemed insufficiently reliable for

very small companies; these are primarily low-degree nodes. Zeeland province is deemed a suitable sub-network on which to focus for its small size and its somewhat lesser integration, geographic and economic, with the rest of the Netherlands. It is also considered an industrial cluster by policy makers and so is expected to have some general internal coherence.

**CBS network reconstruction process.** Dutch national Supply and Use-Tables include aggregated trade flows between industries by product group (Eurostat, 2008a). These are official statistics and extensively validated. CBS has applied systematic inference to dis-aggregate the domestic portion of these flows down to the company level. Customer-supplier ties are inferred based on assumptions made about individual companies with regard to their size, their location, and their role with respect to each product group. Used in this are economic statistics collected via official surveys that include total turnover, geographic location, and industry code according to the Dutch implementation of the Statistical Classification of Economic Activities in the European Community (Eurostat, 2008b, NACE Rev. 2). These surveys are mandatory and use a stratified sampling framework on the number of employees. Note that their aim is to achieve statistically representative, not absolute, accuracy and so inferences are noisy especially for smaller companies. Each product group in the Dutch CPA (2008) becomes a layer in the reconstructed network. Within a layer, companies active in an industry that uses products classified into that product group are assigned a number of suppliers, i.e., an in-degree. It was assumed that most companies deal with few suppliers per product-group; in-degree is kept small. Out-degree, analogously, represents the number of customers for a company active in an industry that supplies products in that product group. Estimations of out-degree were made to reflect empirical distributions with respect to company size known from studies done in Japan (Watanabe et al., 2013). Also meant to be representative, on average, out-degree is noisy for individual companies. Finally, suppliers are matched with customers. Matching was done using limited empirical data on known trade relationships combined with information on geographic distance. The assumption that geographic distance matters is well-supported (Dhyne and Duprez, 2016; Bernard et al., 2019; Carrère et al., 2020). For further details of the CBS network reconstruction process please consult Hooijmaaijers and Buiten (2019).

### 3.1.2. Multi-Layer Complementary Configuration Models

The network theory laid out in section 2 suggests that functional structure might arise in production networks because trade connects companies with complementary roles as customers and suppliers of particular goods. Simulated data is used to assess this explanation. The CBS reconstruction process is already based in matching customers and suppliers where the industry pairing suggests the two companies are compatible in trade, and here we generalize this process in two ways. First, instead of one reconstructed network for Zeeland we consider an ensemble of many possible networks among Zeeland firms with 5+ employees. Second, we leverage the hierarchical structure of the European CPA (2008). The coarser levels of this product

classification are standardized across Europe; commonly studied are the top-level sector products (CPA Level 1) and progressively more detailed CPA Levels 2, 4, and 6. The Dutch national implementation (here, CPA National) is a minor refinement of CPA Level 6.

To generate ensembles of hypothetical company-level production networks we employ a multi-layer directed configuration model. The configuration model is a standard approach for generating ensembles of networks by randomly pairing edge stubs (Newman et al., 2001; Hagberg et al., 2008). We retain the in- and out-degree sequences from each product-layer of the CBS network reconstruction among Zeeland companies with 5+ employees, without self-loops, while drawing many possible wiring instances from the multi-layer configuration model. Our implementation generates a random, directed, multi-network *per product group* in the classification; it then combines the layers into one customer-supplier network. At each CPA Level, the process is repeated multiple times to generate multiple independent realizations from the multi-layer configuration model. The result is a set of networks where ties are trade-compatible relationships under the constraint that trade occurs between companies supplying a product and companies using that same product. Customer-supplier complementarity grows progressively more stringent as we use more detailed product categorizations to define these roles.

Wherever the directed configuration model introduces self-loops, a known artifact of stub-matching (see Newman, 2010), we rewire away the offending links. Specifically, we pair each self-loop with a random other edge and swap their target node (see Hanhijärvi et al., 2009, Figure 1A). On the other hand, when the configuration model generates multi-edges, i.e., edges between the same two nodes, these are allowed to remain; the networks are already multi-layer with many multi-edges.

**Table 3** describes the complementary ensembles, giving the network statistics found across 25 independent realizations at each of the highlighted levels of the European CPA (2008). The resulting multi-layer networks maintain the total number of customer-supplier ties as well as the degree and role of each company in supplying and/or using products with a particular classification. The hierarchical CPA levels introduce variation in the number of product groups considered separately, i.e., the number of layers. The complementarity constraint—that customer-supplier ties may only go from companies who supply to those who use products represented in that layer—is stronger at finer levels. Notice the relative similarity of CPA Level 6 and CPA National. As before, we focus in this work on the simple version of these networks where links correspond to unique trade-compatible relationships. The configuration model introduces randomness in the wiring diagram of the networks and the number of unique trading relationships.

## 3.2. Methods

In this study we analyze the local connectivity structure of the production networks for Zeeland, the whole of the Netherlands, and our complementary ensemble. Section 3.2.1 defines spectral bipartivity, a measure that quantifies the over-representation of even paths in these networks. Functional networks have high

**TABLE 3** | Network statistics for the multi-layer complementary ensembles over 25 realizations, simulating trade-compatible relationships among companies with 5+ employees within Zeeland province.

	Nodes	Multi-edges	Layers	Simple edges	
				Min	Max
<b>ZEELAND, 5+</b>					
CPA Level 1	2497	848,015	18	371,446	372,405
CPA Level 2	2497	848,015	79	369,483	370,190
CPA Level 4	2497	848,015	391	369,303	369,977
CPA Level 6	2497	848,015	614	368,324	368,978
CPA National	2497	848,015	652	368,007	368,813

Layers are designated according to the hierarchical levels of the European CPA (2008) and its Dutch implementation.

values of this measure compared to random expectation, as explained and defined in section 3.2.2. In section 3.2.3 we detail the statistical tools that we use to conduct this comparison. This methodology tested in section 3.2.4 on two public network datasets, producing the expected outcome.

### 3.2.1. Spectral Bipartivity

Spectral methods can be used to summarize the local connectivity structure of a network. The Estrada index is an absolute measure of local connectivity. This measure quantifies the local density of cycles by having closed paths contribute progressively less to the value of the measure, as they take more steps to complete (Estrada, 2000). The value of the Estrada index for a network,  $G$  with  $n$  nodes can be computed as the trace of the matrix exponential of that network's adjacency matrix,  $A$ . Equation (1) gives this definition as well as an alternative formulation where  $\lambda_1 \leq \dots \leq \lambda_n$  are the eigenvalues of  $A$ .

$$EE(G) = \text{tr} \exp(A) = \sum_{j=1}^n e^{\lambda_j} \quad (1)$$

Most relevant to functional networks is a variation of the Estrada index that separates the contribution of even and odd closed paths: spectral bipartivity ( $b_s$ ). This is done using the hyperbolic sine and cosine matrix functions, which add up to the matrix exponential, as applied to a network's adjacency matrix. With proper normalization, spectral bipartivity ranges from 0 when the network is fully complete to 1 when the network is fully bipartite (Estrada and Rodríguez-Velázquez, 2005; Estrada, 2006; Kunegis, 2015; Estrada and Gómez-Gardeñes, 2016). Two-mode networks are the extreme case where the bipartite constraint on link formation entirely disallows odd cycles. Equation (2) defines several equivalent formulations of spectral bipartivity.

$$b_s(G) = \frac{\text{tr} \cosh(A) - \text{tr} \sinh(A)}{\text{tr} \cosh(A) + \text{tr} \sinh(A)} = \frac{\text{tr} \exp(-A)}{\text{tr} \exp(A)} = \frac{\sum_{j=1}^n e^{-\lambda_j}}{\sum_{j=1}^n e^{\lambda_j}} \quad (2)$$

We primarily consider the value of the spectral bipartivity under a logistic transformation, because this metric is restricted in range from 0 to 1 and the Estrada index in the denominator can become quite large. Indeed, it grows exponentially with the square root of the number of edges in certain cases (de la Peña et al., 2007).

Moreover, the logit spectral bipartivity can be readily approximated for networks that fulfill the following two conditions. First, when the spectral bipartivity is very small, i.e.,  $b_s \approx 0$ , the logistic transformation is closely approximated by a log transformation. Second, whenever the most positive and most negative eigenvalues of a network's adjacency matrix are substantially larger in magnitude than their neighboring eigenvalue, i.e.,  $\lambda_1 \ll \lambda_2$  and  $\lambda_n \gg \lambda_{n-1}$ , these eigenvalues will dominate the exponential sums in the numerator and denominator of spectral bipartivity. Equation (3) describes these approximations.

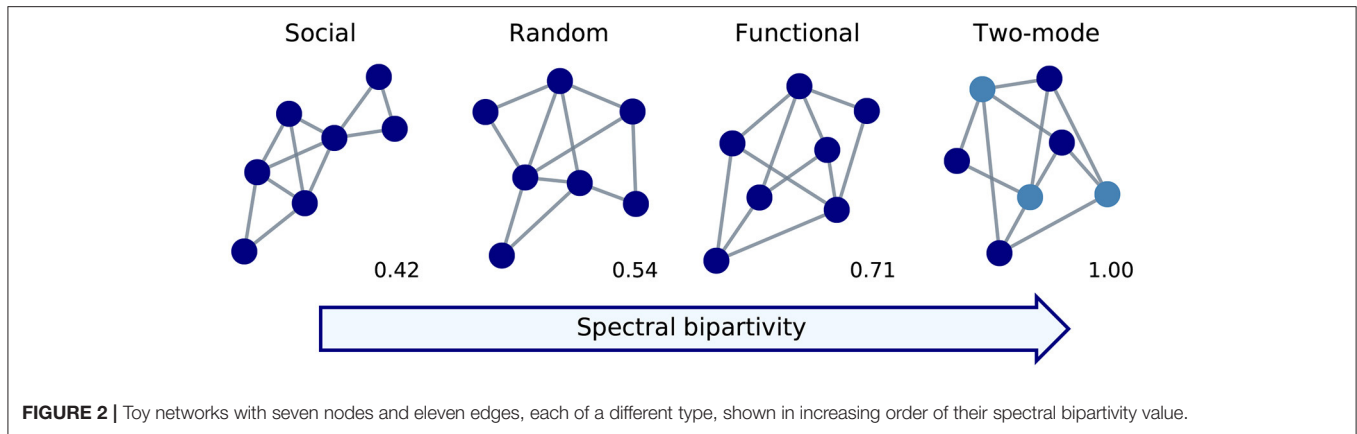
$$\begin{aligned} \text{logit}(b_s) &= \log \frac{b_s}{1 - b_s} = \log b_s - \log(1 - b_s) \approx \log b_s \\ \log(b_s) &= \log \frac{\sum_{j=1}^n e^{-\lambda_j}}{\sum_{j=1}^n e^{\lambda_j}} \approx \log \frac{e^{-\lambda_1}}{e^{\lambda_n}} = -(\lambda_1 + \lambda_n) \end{aligned} \quad (3)$$

### 3.2.2. Identifying Functional Structure

Here we propose that spectral bipartivity can be re-purposed to identify functional structure in networks using a comparison to random expectation. The measure quantifies the over-representation of even paths in the local connectivity structure of a network. Recall that functional networks have especially many squares with even path length, while social networks have especially many triangles with odd path length (see section 2.1). Random networks, with neither social nor functional structure, would produce values of spectral bipartivity that fall in-between those of the other two network types. **Figure 2** gives a toy example of how social, random, functional, and two-mode networks with the same number of nodes and edges are arranged according to their value of spectral bipartivity. Notably, functional networks are more bipartite than random expectation.

The random expectation is found by calculating spectral bipartivity on a set of random networks comparable to our networks of inferred trading relationships. Degree-preserving randomization (Rao et al., 1996; Milo et al., 2004; Gionis et al., 2007) produces random networks that maintain the number of companies, the number of unique inter-company links, and the degree of each company. We use a version of this called random pairwise rewiring, wherein pairs of edges are selected and an end point of each edge are swapped (see Hanhijärvi et al., 2009, Figure 1A). In section 3.1.2, edge swaps were used to remove self-loops.





Here, edge swaps randomize the network. Our implementation guarantees that the randomized network remains simple by following through with an edge swap only so long as it will not introduce self-loops or multi-edges. Randomization continues until  $10 \cdot m$  pairs of links have been rewired, where  $m$  is the number of simple edges.

### 3.2.3. Statistical Test

We use non-parametric statistics to confirm that the networks described in section 3.1 have functional structure. Section 3.2.1 defines spectral bipartivity and section 3.2.2 describes how we generate randomized versions of our networks. The spectral bipartivity computed on these networks are samples drawn from the reference distribution of this measure, as in a Monte Carlo test (see Bartlett, 1963; Besag and Clifford, 1989, 1991, discussion by Barnard, G. A. p. 294). For statistical comparison, we use the two-sample one-tailed Kolmogorov-Smirnov (KS) test (Smirnov, 1948; Dodge, 2008; Virtanen et al., 2020).

The KS statistic quantifies the (lack of) overlap between distributions;  $KS = 0$  indicates identical distributions and  $KS = 1$  indicates non-overlapping distributions, i.e., one is consistently larger than the other. The statistical power of the test is a function of the size of the two distributions and the overlap between them. In analyzing the reconstructed networks, we compare their values of spectral bipartivity against those of their many randomized versions. Finding values consistently and significantly larger-than-random would indicate functional structure. In analyzing our multi-layer complementary configuration models, we consider the difference in logit-transformed spectral bipartivity between each instance and a randomized version of itself. Finding this difference to be consistently and significantly greater than zero would indicate that the model introduces functional structure. Finding one model's distribution of differences to be consistently and significantly larger than another would indicate that it introduces functional structure to a greater extent. Note that comparisons between two distributions have much higher statistical power under non-parametric tests than those between a distribution and a single value.

### 3.2.4. Implementation

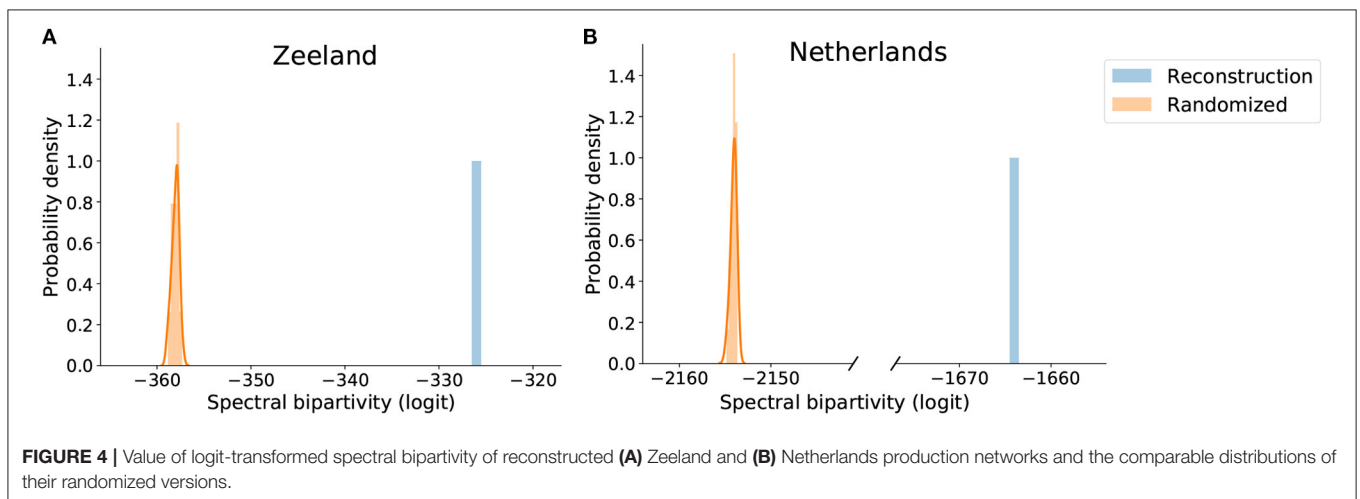
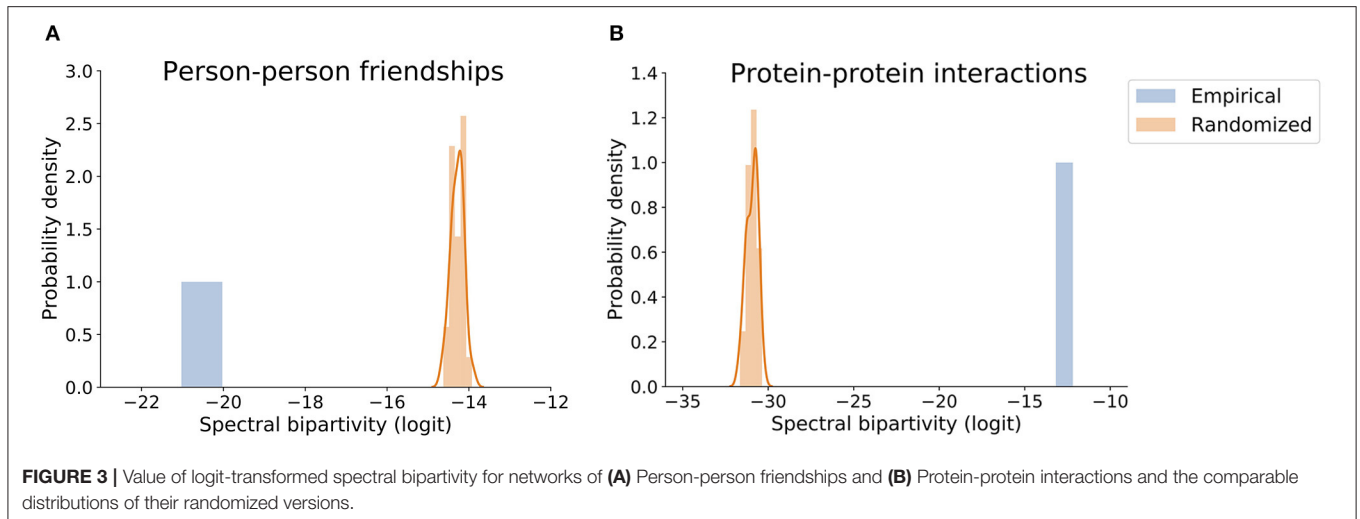
Our implementation of degree-preserving randomization, spectral bipartivity, and its (logit) approximation use `networkx` (Hagberg et al., 2008, v2.5) and `scipy` (Virtanen et al., 2020, v1.5.2). The methodology here described produces the expected result in a demonstration on two public network datasets. The first is a network of Facebook friendships among a group of first-year university students, collected as a part of the Copenhagen Networks Study (Sapiezynski et al., 2019). **Figure 3A** shows its spectral bipartivity is less than random expectation, indicating *social* structure. The second is a network of interactions among proteins in human cells, which was analyzed in Kovács et al. (2019) and is available in that paper's supplementary material. **Figure 3B** shows its spectral bipartivity is greater than random expectation, indicating *functional* structure. The code to produce these figures is made available at <https://github.com/carolinamattsson/local-connectivity-structure>.

## 4. RESULTS

In this section we present our analysis of local connectivity structure in company-level production networks. The reconstructed Zeeland and Netherlands networks show functional structure (section 4.1) and this arises, at least in part, due to the complementary nature of customer-supplier ties (section 4.2).

### 4.1. Local Connectivity Structure

We find functional structure in the reconstructed company-level production networks. The spectral bipartivity of the network of (inferred) trade relationships in Zeeland is substantially larger than random expectation; statistical comparison yields a one-tailed Kolmogorov-Smirnov statistic of 1.0 ( $N_1 = 1$ ,  $N_2 = 1000$ ,  $p < 0.001$ ). **Figure 4A** plots the logit-transformed value of spectral bipartivity for the Zeeland production network against the distribution of values over 25 randomized instances. The network falls well above random expectation and can be said to have functional structure. The value of spectral bipartivity (in the absolute sense) is small, at  $1.6 \cdot 10^{-142}$ , indicating a



lack of bipartite structure and that the network is definitely not two-mode.

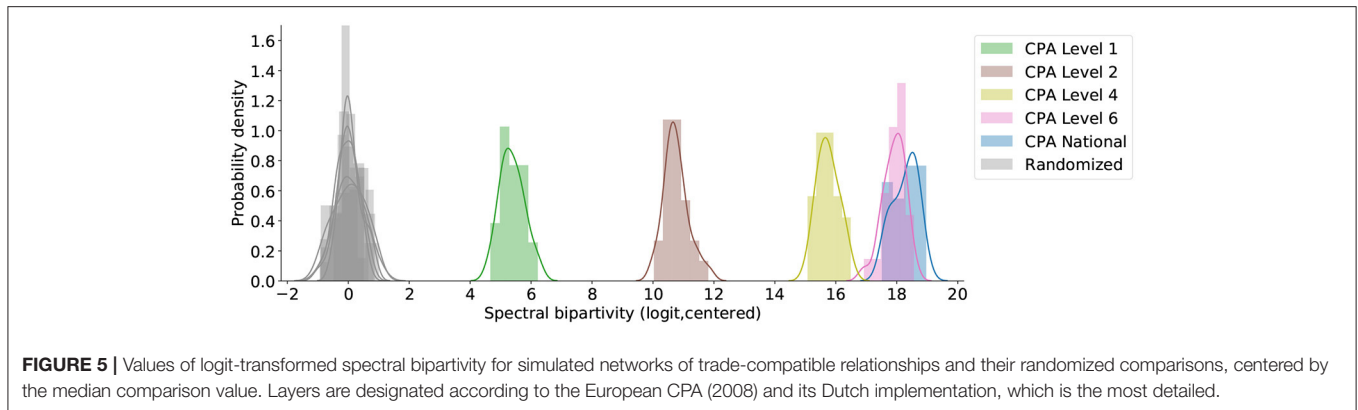
Using the approximation described in Equation (3), we extend this result to the company-level production network for the whole of the Netherlands (**Figure 4B**). This network, consisting of 50,930,077 inferred trade relationships among 102,461 companies with 5+ employees, also has functional structure. Its logit spectral bipartivity is significantly higher-than-random ( $KS = 1.0$ ,  $N_1 = 1$ ,  $N_2 = 25$ ,  $p < 0.04$ ). Also using the eigenvalue approximation, we confirm that our result holds for the reconstructed network of inferred trade relationships among all 18,398 companies in Zeeland province, including those with fewer than 5 employees. Again this network is significantly more bipartite than random expectation, indicating functional structure ( $KS = 1.0$ ,  $N_1 = 1$ ,  $N_2 = 25$ ,  $p < 0.04$ ).

In interpreting these results, recall that the comparison to random expectation is key to our re-purposing of spectral bipartivity for analyzing local connectivity structure. Values of spectral bipartivity, in the absolute sense, reflect also other network-structural features. In particular, this measure is defined as a ratio and the denominator of this ratio can be strongly

affected by the number of edges in the network (see section 3.2.1). The Zeeland network has a higher value of spectral bipartivity than the Netherlands network because the Netherlands network is substantially larger in size. Randomization preserves the number of edges (and the degree sequence) so it is serving as a way to “center” the scale such that remaining differences reflect local connectivity structure; this lets us identify functional structure. Developing measures that would allow for comparison across networks substantially different in size is a promising area for future work.

## 4.2. Customer-Supplier Complementarity

We find evidence that functional structure in production networks is driven by the principle of node complementarity, i.e., the practical fact that trade relationships imply the exchange of some product from a supplier to a customer. Our multi-layer configuration models allow customer-supplier ties only between pairs of companies compatible in trade; a likely supplier and a likely user of products with a particular classification (see section 3.1.2). Each of the



resulting networks reflect trade-compatible relationships between companies with complementary roles in this network. This constraint gets progressively more stringent as the product classification that defines the roles gets more detailed (see **Table 3**).

Already at the top, sector-level, classification (CPA Level 1) customer-supplier complementarity introduces some functional structure. **Figure 5** compares 25 instances from our configuration model at CPA Level 1 to their randomized versions, where the relationships are no longer necessarily trade-compatible. The logit-transformed spectral bipartivity is consistently larger-than-random where customer-supplier complementarity has been introduced at CPA Level 1 ( $KS = 1.0$ ,  $N_1 = 1$ ,  $N_2 = 25$ ,  $p < 0.04$ ). Stricter complementarity constraints at the more detailed CPA Levels 2, 4, and 6 each introduce consistently and significantly more functional structure ( $KS = 1.0$ ,  $N_1 = 25$ ,  $N_2 = 25$ ,  $p < 10^{-14}$ ). On the other hand, and as expected, the small refinement between CPA Level 6 and the CPA National product classification produces a minor increase in the separation between model and random expectation; it is statistically, but not substantively, significant ( $KS = 0.44$ ,  $N_1 = 25$ ,  $N_2 = 25$ ,  $p < 0.01$ ).

In interpreting this finding, recall that these are simulated networks whose links are sampled from among the many possible links between trade-compatible companies. The roles of individual companies are defined with respect to standardized industry and product categorizations used in the collection of official economic statistics throughout Europe (Eurostat, 2008b). These categories are broad relative to the number of differentiated products traded within the Netherlands and so trade relationships in our networks reflect potential trade in a number of relevant products. This means that the constraint we impose in generating our networks is fairly blunt, even at CPA National and especially at CPA Level 1. In this way, our generalization of the CBS network reconstruction process serves as an important robustness check on the results in section 4.1. At the same time, functional structure could come about via various channels. Most trivially, two suppliers and two users of the same product could form a closed square within one layer due to redundant purchasing patterns. This is unlikely to be what we are picking up on as it is assumed that the vast

majority of companies deal with only one or a few suppliers per product-group (see section 3.1.1). More common, here, would be the situation where two suppliers of two different products sell those products to the same two customers; this would form a closed square when the two layers are collapsed into a simple network of trade relationships.

## 5. DISCUSSION

This paper has advanced the hypothesis that company-level production networks are so-called “functional” networks, with a distinctive local connectivity structure first identified in network biology (Kovács et al., 2019; Kitsak, 2020). This hypothesis was made concrete through a discussion on local connectivity structures and the re-interpretation of previous empirical findings through this lens (Fujiwara and Aoyama, 2010; Ohnishi et al., 2010). We then explored this hypothesis, directly, using a re-purposed network metric (Estrada and Gómez-Gardeñes, 2016) and an existing dataset of inferred customer-supplier ties produced by Statistics Netherlands (CBS) (Hooijmaaijers and Buiten, 2019). This methodology identifies functional structure in reconstructed production networks representing trade among companies in Zeeland province and the whole of the Netherlands. Our generalization of the CBS network reconstruction process, using multi-layer configuration models, then illustrates that customer-supplier complementarity is key to the emergence of functional structure in company-level production networks.

In interpreting these findings, our company-level production networks should be understood as a selection of likely trade relationships where the industry pairing implies the two companies are compatible in trade. Much nuance is avoided in that the industry and product categorizations used in the generation of these networks are standardized and broad; inferred ties reflect potential trade in a number of relevant products. However, as noted in section 3.1.1, inferences about individual companies are noisy as the underlying data are collected for statistical purposes. For this reason, we limit our network analysis to characterizing whole-network patterns in local connectivity structure. To illustrate functional structure in production networks at a more intuitive scale, section 2.2

highlights prior work that finds closed squares and interpretable functional modules in empirical data; it is reassuring that this fully supports our conclusions. The assumptions made in the CBS network reconstruction process when matching customers with suppliers could more directly affect the local connectivity structure of our networks. On this point, it is reassuring that functional structure persists also under our generalized reconstruction process and using less detailed product classifications. Exploring a wider range of data-informed network generation processes would allow more specific hypotheses to be tested and is a promising direction for future work.

Our findings have practical implications for the analysis of company-level production networks and wider implications for the study of economic systems. Practically speaking, our work helps narrow down the set of network analysis techniques relevant for use with networks representing trade relationships among companies, be they empirical, reconstructed, or modeled. To be interpretable, these should conform to the logic of functional networks: companies trade with complementary others and it is close competitors—companies with many shared customers and suppliers—who are especially similar. For the problem of link prediction, for instance, heuristics that close squares (L3) are expected to be especially accurate (Kovács et al., 2019). On the other hand, there is little reason to expect high levels of clustering or high levels of reciprocity between customers and suppliers. Intuitions and techniques developed within network biology for use with PPI networks are likely to be especially applicable (Barabási and Oltvai, 2004; Kitsak, 2020). For instance, identifiable meso-scale structures can be interpreted more readily as “functional modules” (Chen and Yuan, 2006) than as “communities” (Blondel et al., 2008).

Speaking more broadly, structural features of functional networks can deepen our understanding of short-term dynamics on production networks. Several ways economic systems can fail are already studied using detailed simulations over empirical company-level production networks (Hazama and Uesugi, 2017; Inoue and Todo, 2020). From such studies, supply disruptions are known to compound when also the competitors of affected companies become affected via shared customers and suppliers. This phenomenon is related to the prominence of closed squares in the local connectivity structure. An anecdotal example is that during the 2008 financial crisis the CEO of Ford, Alan Mulally, gave testimony in favor of US Government support for General Motors and Chrysler. Mulally argued that the demise of his competitors would imperil shared suppliers of highly specialized auto parts (Klier and Rubenstein, 2013, p. 145). Note that the Japanese automotive industry, also, is especially locally bipartite with many closed squares (Fujiwara and Aoyama, 2010, p. 570). In the demand direction, one might consider the analogous phenomenon where Chinese e-commerce platforms selling personal protective equipment to many of the same customers all experienced shortages in sourcing from many of the same manufacturers following reports of a deadly

outbreak of infectious disease in Wuhan, China in January, 2020 (McMorrow and Liu, 2020).

Local connectivity structure can also help in the study of routine function and growth of production networks happening over longer timescales. Identifying functional structure in these networks has given us precise terminology with which to describe their defining features and the possible link-level mechanisms at play. Latent geometry lends us the term “node complementarity” for the notion that two companies are more likely to interact if one is similar to the other’s other trading partners (Kitsak, 2020); “closed squares” feature prominently (Kovács et al., 2019). It may be especially relevant to study the impact of indirect connections between competitors (via shared customers and suppliers) on the establishment of price and trust across trade relationships, which are thought to be key to productivity improvements at the company level (Uzzi, 1997; Cardoso et al., 2019). Network biology lends us the term “functional module” for meso-scale structures that let the network perform more complex higher-level functions (Barabási and Oltvai, 2004; Chen and Yuan, 2006; Ghiassian et al., 2015). Interpretable modules have been described in detail on empirical company-level production networks (Fujiwara and Aoyama, 2010, p. 570) and may be related to the development of new capabilities in local economies (Dawley, 2014; Boschma et al., 2017). Studying these specific local structural features of production networks could further our understanding of their routine function, failure, and growth.

## DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: Statistics Netherlands (CBS). Requests to access these datasets should be directed to GB, g.buiten@cbs.nl.

## AUTHOR CONTRIBUTIONS

EH, CM, and FT developed the theory. CD, CM, and FT determined the method. GB and CM performed the analysis. AF and PS managed the project. CM wrote the manuscript. All authors discussed the results and edited the final manuscript.

## FUNDING

This work was supported by the Ministry for Economic Affairs and Climate, The Netherlands.

## ACKNOWLEDGMENTS

The authors acknowledge Joyce Ten Holter and Gideon Mooijen for valuable assistance. We thank Leo Torres, István Kovács, and Maksim Kitsak for discussion.

## REFERENCES

- Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica* 80, 1977–2016. doi: 10.3982/ECTA9623
- Arthur, W. B. (2021). Foundations of complexity economics. *Nat. Rev. Phys.* 3, 136–145. doi: 10.1038/s42254-020-00273-3
- Asikainen, A., Iñiguez, G., Ureña-Carrión, J., Kaski, K., and Kivelä, M. (2020). Cumulative effects of triadic closure and homophily in social networks. *Sci. Adv.* 6:eaax7310. doi: 10.1126/sciadv.aax7310
- Barabási, A. L., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113. doi: 10.1038/nrg1272
- Barabási, A. L., and Pósfai, M. (2016). "Chapter 7: degree correlations," in *Network Science* (Cambridge: Cambridge University Press), 232–296.
- Barber, M. J. (2007). Modularity and community detection in bipartite networks. *Phys. Rev. E* 76:066102. doi: 10.1103/PhysRevE.76.066102
- Bartlett, M. S. (1963). The spectral analysis of point processes. *J. R. Statist. Soc. B Methodol.* 25, 264–296. doi: 10.1111/j.2517-6161.1963.tb00508.x
- Berardo, R. (2014). Bridging and bonding capital in two-mode collaboration networks. *Policy Stud. J.* 42, 197–225. doi: 10.1111/psj.12056
- Bernard, A. B., Moxnes, A., and Saito, Y. U. (2019). Production networks, geography, and firm performance. *J. Polit. Econ.* 127, 639–688. doi: 10.1086/700764
- Besag, J., and Clifford, P. (1989). Generalized Monte Carlo significance tests. *Biometrika* 76, 633–642. doi: 10.1093/biomet/76.4.633
- Besag, J., and Clifford, P. (1991). Sequential Monte Carlo *p*-values. *Biometrika* 78, 301–304. doi: 10.1093/biomet/78.2.301
- Blondel, V. D., Guillaume, J. L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Statist. Mech. Theory Exp.* 2008:P10008. doi: 10.1088/1742-5468/2008/10/P10008
- Bollobás, B. (2001). *Random Graphs*. Cambridge: Cambridge University Press. Available online at: <https://www.cambridge.org/core/books/random-graphs/E21023008001CFA182CE666F5028489F>
- Borgatti, S. P., and Everett, M. G. (1997). Network analysis of 2-mode data. *Soc. Netw.* 19, 243–269. doi: 10.1016/S0378-8733(96)00301-2
- Boschma, R., Coenen, L., Frenken, K., and Truffer, B. (2017). Towards a theory of regional diversification: combining insights from evolutionary economic geography and transition studies. *Reg. Stud.* 51, 31–45. doi: 10.1080/00343404.2016.1258460
- Bounova, G., and de Weck, O. (2012). Overview of metrics and their correlation patterns for multiple-metric topology analysis on heterogeneous graph ensembles. *Phys. Rev. E* 85:016117. doi: 10.1103/PhysRevE.85.016117
- Burkholz, R. (2016). *Systemic Risk: From Generic Models to Food Trade Networks*. Technical report, ETH Zurich.
- Cardoso, F. M., Gracia-Lazaro, C., Moisan, F., Goyal, S., Sanchez, A., and Moreno, Y. (2019). Trading in complex networks. *arXiv* 1906.01531.
- Carrère, C., Mrázová, M., and Neary, J. P. (2020). Gravity without apology: the science of elasticities, distance and trade. *Econ. J.* 130, 880–910. doi: 10.1093/ej/ueaa034
- Carvalho, V. M., and Tahbaz-Salehi, A. (2018). *Production Networks: A Primer*. Technical Report 1856, Faculty of Economics, University of Cambridge.
- Chen, J., and Yuan, B. (2006). Detecting functional modules in the yeast protein–protein interaction network. *Bioinformatics* 22, 2283–2290. doi: 10.1093/bioinformatics/btl370
- Coenen, L., Raven, R., and Verbong, G. (2010). Local niche experimentation in energy transitions: a theoretical and empirical exploration of proximity advantages and disadvantages. *Technol. Soc.* 32, 295–302. doi: 10.1016/j.techsoc.2010.10.006
- Colomer-de Simón, P., Serrano, M. A., Beiró, M. G., Alvarez-Hamelin, J. I., and Boguñá, M. (2013). Deciphering the global organization of clustering in real complex networks. *Sci. Rep.* 3:2517. doi: 10.1038/srep02517
- Costa, L. F., Rodrigues, F. A., Travieso, G., and Boas, P. R. V. (2007). Characterization of complex networks: a survey of measurements. *Adv. Phys.* 56, 167–242. doi: 10.1080/00018730601170527
- Davis, A., Gardner, B. B., and Gardner, M. R. (2009). *Deep South: A Social Anthropological Study of Caste and Class*. Columbia, SC: University of South Carolina Press.
- Dawley, S. (2014). Creating new paths? Offshore wind, policy activism, and peripheral region development. *Econ. Geogr.* 90, 91–112. doi: 10.1111/ecge.12028
- de la Peña, J. A., Gutman, I., and Rada, J. (2007). Estimating the Estrada index. *Linear Algebra Appl.* 427, 70–76. doi: 10.1016/j.laa.2007.06.020
- Dhyne, E., and Duprez, C. (2016). Three regions, three economies? *Econ. Rev.* 2016, 59–73. Available online at: <https://www.nbb.be/en/articles/three-regions-three-economies>
- Dodge, Y. (Eds.). (2008). "Kolmogorov–Smirnov test," in *The Concise Encyclopedia of Statistics* (New York, NY: Springer), 283–287.
- Estrada, E. (2000). Characterization of 3D molecular structure. *Chem. Phys. Lett.* 319, 713–718. doi: 10.1016/S0009-2614(00)00158-5
- Estrada, E. (2006). Protein bipartivity and essentiality in the yeast protein–protein interaction network. *J. Proteome Res.* 5, 2177–2184. doi: 10.1021/pr060106e
- Estrada, E., and Gómez-Gardeñes, J. (2016). Network bipartivity and the transportation efficiency of European passenger airlines. *Phys. D Nonlin. Phenom.* 323–324, 57–63. doi: 10.1016/j.physd.2015.10.020
- Estrada, E., and Rodríguez-Velázquez, J. A. (2005). Spectral measures of bipartivity in complex networks. *Phys. Rev. E* 72:046105. doi: 10.1103/PhysRevE.72.046105
- Eurostat (2008a). *Eurostat Manual of Supply, Use and Input-Output Tables, 2008 Edn*. Luxembourg: Office for Official Publications of the European Communities.
- Eurostat (2008b). *NACE rev. 2*. Luxembourg: Office for Official Publications of the European Communities.
- Fujiwara, Y., and Aoyama, H. (2010). Large-scale structure of a nation-wide production network. *Eur. Phys. J. B* 77, 565–580. doi: 10.1140/epjb/e2010-00275-2
- Ghasemian, A., Hosseinmardi, H., Galstyan, A., Airolidi, E. M., and Clauset, A. (2020). Stacking models for nearly optimal link prediction in complex networks. *Proc. Natl. Acad. Sci. U.S.A.* 117, 23393–23400. doi: 10.1073/pnas.1914950117
- Ghiassian, S. D., Menche, J., and Barabási, A. L. (2015). A DIsease MOdule Detection (DIAMOnD) algorithm derived from a systematic analysis of connectivity patterns of disease proteins in the human interactome. *PLoS Comput. Biol.* 11:e1004120. doi: 10.1371/journal.pcbi.1004120
- Gionis, A., Mannila, H., Mielikäinen, T., and Tsaparas, P. (2007). Assessing data mining results via swap randomization. *ACM Trans. Knowl. Discov. Data* 1:14. doi: 10.1145/1297332.1297338
- Hagberg, A. A., Schult, D. A., and Swart, P. J. (2008). "Exploring network structure, dynamics, and function using networkX," in *Proceedings of the 7th Python in Science Conference* (Pasadena, CA), 5.
- Hanhijärvi, S., Garriga, G. C., and Puolamäki, K. (2009). "Randomization techniques for graphs," in *Proceedings of the 2009 SIAM International Conference on Data Mining, Proceedings* (Society for Industrial and Applied Mathematics), 780–791. doi: 10.1137/1.9781611972795.67
- Hausmann, R., and Hidalgo, C. A. (2013). *How Will the Netherlands Earn Its Income 20 Years From Now*. The Hague: The Netherlands Scientific Council for Government Policy. Available online at: <https://english.wrr.nl/publications/publications/2013/11/04/how-will-the-netherlands-earn-its-income-20-years-from-now>
- Hazama, M., and Uesugi, I. (2017). Measuring the systemic risk in interfirm transaction networks. *J. Econ. Behav. Organ.* 137, 259–281. doi: 10.1016/j.jebo.2017.02.009
- Holland, P. W., Laskey, K. B., and Leinhardt, S. (1983). Stochastic blockmodels: first steps. *Soc. Netw.* 5, 109–137. doi: 10.1016/0378-8733(83)90021-7
- Holme, P., Liljeros, F., Edling, C. R., and Kim, B. J. (2003). Network bipartivity. *Phys. Rev. E* 68:056107. doi: 10.1103/PhysRevE.68.056107
- Hooijmaaijers, S., and Buiten, G. (2019). *A Methodology for Estimating the Dutch Interfirm Trade Network, Including a Breakdown by Commodity*. Technical report, Statistics Netherlands.
- Inoue, H., and Todo, Y. (2019). Firm-level propagation of shocks through supply-chain networks. *Nat. Sustain.* 2, 841–847. doi: 10.1038/s41893-019-0351-x
- Inoue, H., and Todo, Y. (2020). The propagation of the economic impact through supply chains: the case of a mega-city lockdown against the spread of COVID-19. *arXiv* 2003.14002. doi: 10.2139/ssrn.3564898
- Jamakov, A., Mahadevan, P., Vahdat, A., Boguna, M., and Krioukov, D. (2015). How small are building blocks of complex networks. *arXiv* 0908.1143.

- Jamakovic, A., and Uhlig, S. (2008). On the relationships between topological measures in real-world networks. *Netw. Heterogen. Media* 3:345. doi: 10.3934/nhm.2008.3.345
- Johnson, S., Torres, J. J., Marro, J., and Muñoz, M. A. (2010). Entropic origin of disassortativity in complex networks. *Phys. Rev. Lett.* 104:108702. doi: 10.1103/PhysRevLett.104.108702
- Karrer, B., and Newman, M. E. J. (2011). Stochastic blockmodels and community structure in networks. *Phys. Rev. E* 83:016107. doi: 10.1103/PhysRevE.83.016107
- Kitsak, M. (2020). Latent geometry for complementarity-driven networks. *arXiv* 2003.06665.
- Klier, T., and Rubenstein, J. M. (2013). Restructuring of the U.S. auto industry in the 2008–2009 recession. *Econ. Dev. Q.* 27, 144–159. doi: 10.1177/0891242413481243
- Kovács, I. A., Luck, K., Spirohn, K., Wang, Y., Pollis, C., Schlabach, S., et al. (2019). Network-based prediction of protein interactions. *Nat. Commun.* 10:1240. doi: 10.1038/s41467-019-09177-y
- Kunegis, J. (2015). Exploiting the structure of bipartite graphs for algebraic and spectral graph theory applications. *Internet Math.* 11, 201–321. doi: 10.1080/15427951.2014.958250
- Magerman, G., Dhyne, E., and Rubinova, S. (2015). *The Belgian Production Network 2002–2012*. Technical Report 288, National Bank of Belgium.
- Masuda, N., Sakaki, M., Ezaki, T., and Watanabe, T. (2018). Clustering coefficients for correlation networks. *Front. Neuroinform.* 12:7. doi: 10.3389/fninf.2018.00007
- McMorrow, R., and Liu, N. (2020). *Chinese Ecommerce Groups Struggle to Meet Demand for Coronavirus Supplies*. London: Financial Times.
- McNerney, J., Savoie, C., Caravelli, F., and Farmer, J. D. (2018). How production networks amplify economic growth. *arXiv* 1810.07774.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: homophily in social networks. *Annu. Rev. Sociol.* 27, 415–444. doi: 10.1146/annurev.soc.27.1.415
- Miller, R. E., and Temurshoev, U. (2017). Output upstreamness and input downstreamness of industries/countries in world production. *Int. Reg. Sci. Rev.* 40, 443–475. doi: 10.1177/0160017615608095
- Milo, R., Kashtan, N., Itzkovitz, S., Newman, M. E. J., and Alon, U. (2004). On the uniform generation of random graphs with prescribed degree sequences. *arXiv cond-mat/0312028*.
- Mizruchi, M. S. (1996). What do interlocks do? An analysis, critique, and assessment of research on interlocking directorates. *Annu. Rev. Sociol.* 22, 271–298. doi: 10.1146/annurev.soc.22.1.271
- Newman, M. E. J. (2003a). Mixing patterns in networks. *Phys. Rev. E* 67:026126. doi: 10.1103/PhysRevE.67.026126
- Newman, M. E. J. (2003b). The structure and function of complex networks. *SIAM Rev.* 45, 167–256. doi: 10.1137/S003614450342480
- Newman, M. E. J. (2006). Modularity and community structure in networks. *Proc. Natl. Acad. Sci. U.S.A.* 103, 8577–8582. doi: 10.1073/pnas.0601602103
- Newman, M. E. J. (2010). *Random Graphs With General Degree Distributions*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/978019206650.003.0013
- Newman, M. E. J., and Park, J. (2003). Why social networks are different from other types of networks. *Phys. Rev. E* 68:036122. doi: 10.1103/PhysRevE.68.036122
- Newman, M. E. J., Strogatz, S. H., and Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* 64:026118. doi: 10.1103/PhysRevE.64.026118
- Ohnishi, T., Takayasu, H., and Takayasu, M. (2010). Network motifs in an inter-firm network. *J. Econ. Interact. Coord.* 5, 171–180. doi: 10.1007/s11403-010-0066-6
- Opsahl, T. (2013). Triadic closure in two-mode networks: redefining the global and local clustering coefficients. *Soc. Netw.* 35, 159–167. doi: 10.1016/j.socnet.2011.07.001
- Pastor-Satorras, R., Vázquez, A., and Vespignani, A. (2001). Dynamical and correlation properties of the internet. *Phys. Rev. Lett.* 87:258701. doi: 10.1103/PhysRevLett.87.258701
- Rao, A. R., Jana, R., and Bandyopadhyay, S. (1996). A Markov chain Monte Carlo method for generating random (0 1)-matrices with given marginals. *Sankhy* 58, 225–242.
- Rivera, M. T., Soderstrom, S. B., and Uzzi, B. (2010). Dynamics of Dyads in social networks: assortative, relational, and proximity mechanisms. *Annu. Rev. Sociol.* 36, 91–115. doi: 10.1146/annurev.soc.34.040507.134743
- Sapiezynski, P., Stopczynski, A., Lassen, D. D., and Lehmann, S. (2019). Interaction data from the Copenhagen Networks Study. *Sci. Data* 6:315. doi: 10.1038/s41597-019-0325-x
- Seierstad, C., and Opsahl, T. (2011). For the few not the many? The effects of affirmative action on presence, prominence, and social capital of women directors in Norway. *Scand. J. Manage.* 27, 44–54. doi: 10.1016/j.scaman.2010.10.002
- Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *Ann. Math. Stat.* 19, 279–281. doi: 10.1214/aoms/1177730256
- Statline (2012). *Bevolkingontwikkeling*. Den Haag: National Statistics, Statistics Netherlands.
- Takes, F. W., and Heemskerck, E. M. (2016). Centrality in the global network of corporate control. *Soc. Netw. Anal. Mining* 6:97. doi: 10.1007/s13278-016-0402-5
- Uzzi, B. (1997). social structure and competition in interfirm networks: the paradox of embeddedness. *Admin. Sci. Q.* 42, 35–67. doi: 10.2307/2393808
- Valeeva, D., Heemskerck, E. M., and Takes, F. W. (2020). The duality of firms and directors in board interlock networks: a relational event modeling approach. *Soc. Netw.* 62, 68–79. doi: 10.1016/j.socnet.2020.02.009
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* 17, 261–272. doi: 10.1038/s41592-019-0686-2
- Watanabe, H., Takayasu, H., and Takayasu, M. (2013). Relations between allometric scalings and fluctuations in complex systems: the case of Japanese firms. *Phys. A Statist. Mech. Appl.* 392, 741–756. doi: 10.1016/j.physa.2012.10.020
- Zhou, T., Ren, J., Medo, M., and Zhang, Y. C. (2007). Bipartite network projection and personal recommendation. *Phys. Rev. E* 76:046115. doi: 10.1103/PhysRevE.76.046115

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Mattsson, Takes, Heemskerck, Diks, Buiten, Faber and Sloot. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.