Discussion Paper

Triadic Pattern Detection on Inter-Industry Production Networks: A Multi-Commodity Analysis

M. Di Vece
D. Garlaschelli
F.P. Pijpers

November 2024

Contents

Triadic motifs are the smallest interconnected building blocks of complex networks, such as production networks. They can be detected as over-occurrences with respect to null models that only consider pairwise interactions. Recently, there has been growing interest in the role of triadic motifs in the propagation of economic shocks. However, their characterization at the level of individual commodities is still poorly understood.

To address this gap, we analyze both binary and weighted triadic motifs in the Dutch inter-industry production network disaggregated at the level of 187 commodity groups. We introduce appropriate null models that filter out node heterogeneity and the strong effects of link reciprocity. We find that, while the aggregate network that overlays all products is characterized by a multitude of triadic motifs, most single-product layers feature no significant motif, and roughly 85% of the layers feature only two motifs or less.

This result has several important policy implications. First, it suggests that the propagation of economic shocks through production networks is likely to be complex and heterogeneous, as it will depend on the specific triadic motifs that are present in each commodity layer. Second, it implies that statistical bureaus can identify fine-grained information about the structural relationships between different commodities by analyzing triadic motifs at the disaggregated level. This information can be used to develop more targeted and effective policy interventions.

Overall, the analysis of triadic motifs in production networks has the potential to provide policymakers with valuable insights into the structure and dynamics of these networks. This information can be used to develop more targeted and effective policies to support economic growth and resilience.

Keywords: production network, pattern detection,; network motifs

1 Introduction

1.1 Background

Over the past decade, research into the complex network of relationships between statistical units is a burgeoning field for national statistical institutes. Statistics Netherlands (CBS) is playing a pioneering role in this due to the wealth of integral, register-based, data that is available. There are now micro-level network data availlable for research, not only for researchers within CBS but also for academic research through the standard protocols for secure access to such data. For relationships between people, at least those relationships that are registered for administrative purposes, there are now datasets for a number of years so that the dynamics of social networks can start to be studied. For business interactions and firm-to-firm trade there is a parallel effort in reproducing the structure of economic and financial networks. While there is more limited information available to Statistics Netherlands on firm-to-firm transactions, the available data does allow reconstruction and it has to that end developed its own method to do so, cf. (Hooijmaaijers and Buiten 2019).

While aggregated information about single firms is contained in most National Statistical Institutes' repositories, reliable data on input/output relationships are available only for a small number of countries. For instance, the Compustat dataset contains the major customers of the publicly listed firms in the USA (Atalay et al. 2011). The FactSet Revere dataset contains major customers of publicly listed firms at a global level, with a focus on the USA, Europe, and Asia (König et al. 2022). Two datasets are commercially available in Japan, namely, the dataset collected by Tokyo Shoko Research Ltd. (TSR) (Carvalho et al. 2020) and the one collected by Teikoku DataBank Inc. (TDB) (Mizuno et al. 2014). They are characterized by a high coverage of Japanese firms but with a limited amount of commercial partners. Other domestic datasets contain transaction values among VAT-liable firms: this is the case for countries such as Brazil (Mungo et al. 2023), Belgium (Dhyne et al. 2015), Hungary (Diem et al. 2022), Ecuador (Bacilieri et al. 2023), Kenya (Chacha et al. 2022), Turkey (Demir et al. 2022), Spain (Peydro et al. 2020), Rwanda and Uganda (Spray and Wolf 2018), West Bengal (Kumar et al. 2021); or contain transaction values among the totality of registered domestic firms such as in the case of Dominican Republic (Cardoza et al. 2020) and Costa Rica (Alfaro-Ureña et al. 2018).

Recently, Statistics Netherlands (CBS) produced two multi-layer production network datasets for domestic intermediate trade of Dutch firms for 2012 (Hooijmaaijers and Buiten 2019) and 2018 (Buiten et al. 2021), with each layer corresponding to a different product exchanged by a firm for its own production process. The presence of product granularity makes it an invaluable source for the analysis of commodity-specific structural patterns. The importance of such patterns for structural analysis of the (Dutch) economy is underlined by e.g. Chong et al. 2019, and (Vries et al. 2021), which build on extracting such knowledge from the aggregates recorded in supply-use tables.

The 2012 dataset has been recently used to prove the complementarity structure of production networks (Mattsson et al. 2021) by inspecting the number of cycles of order 3 and 4 compared to a null model taking into consideration the in-degree and out-degree distributions. Firms are matched according to a deterministic procedure, which is shown to decrease the dataset quality by inducing a bias in the network density and the degree distribution, as proved in (Rachkov et al. 2021) for a sample of known links of the production network collected by Dun & Bradstreet. We use the improved version for 2018 and construct an inter-industry network.

This discussion paper is a companion to Di Vece et al. 2024, in which the details of the methodology are explained. Here, the emphasis is on the results for individual industry groups at various levels of aggregation.

1.2 Relevance

In production networks, user firms connect to supplier firms to buy goods for their own production. Customer-Supplier relationships are, hence, characterized by an intrinsic product granularity. Any given firm can in principle operate more than one production pipeline and is capable of supplying multiple products. For example, a well-known multinational telecommunications company also sells household appliances, and multinational companies known for multimedia are also suppliers or re-sellers of a large number of different products. The aggregation from product-level to firm-level relationships is non-trivial and can produce structural changes in the resulting inferred network.

The heterogeneity encoded in the production network structure plays an essential role in structuring national economies and controlling their dynamics, whether that be the economic response to exogenous (international) factors or endogenously created volatility. Studies of the former are e.g. on economic growth (McNerney et al. 2022) and in the propagation of shocks (Acemoglu et al. 2012), (Carvalho and Tahbaz-Salehi 2019) related to exogenous events, or endogenous events such as the 2008 financial crisis (Maluck and Donner 2015; Wang et al. 2022).

In this study, we focus on triadic motifs and anti-motifs that are over-occurrences and under-occurrences of different patterns of directed triadic connections, respectively. Triadic and tetradic connections are known as the building blocks of complex networks (Milo et al. 2002), playing the role of homophily-driven connections in social networks (Asikainen et al. 2020), or complementarity-driven structures in production networks (Ohnishi et al. 2010; Mattsson et al. 2021). It has been proven that for the majority of (available) real-world networks, the triadic structure is maximally random (Colomer-de-Simón et al. 2013) and by fixing it their global structure is statistically determined (Jamakovic et al. 2009). In contrast, research on weighted motifs and anti-motifs is still underdeveloped. To our knowledge, only one study involves trade volumes circulating on triadic subgraphs, using a probabilistic model based on random walks on the WTW (Picciolo et al. 2022).

Motif detection strictly depends not only on the properties of the real network but also on the randomization method used for the computation of random expectations. In Network Science literature, various methods have been advanced for network randomization, primarily edge-stub methods, edge-swapping methods, and Maximum-Entropy methods, we focus on the latter. Randomization methods based on Entropy Maximization (Jaynes 1957a; Jaynes 1957b; Jaynes 1982) build Graph Probability distributions that are maximally random by construction. Available global or node-specific data are encoded as constraints in the optimization procedure, and their corresponding Lagrange Multipliers are computed by Maximum Likelihood Estimation (MLE) (Garlaschelli and Loffredo 2008). This theoretical framework has been proven to successfully reconstruct economic and financial systems (Bardoscia et al. 2021; Cimini et al. 2019; Cimini et al. 2021; Squartini and Garlaschelli 2017), accurately predicting both the topology and the weights of the WTW (Garlaschelli and Loffredo 2004; Squartini et al. 2011a; Squartini et al. 2011b).

Two studies using Maximum-Entropy modeling are especially worthy of note for motif detection: a theoretical study where the authors develop null models for triadic motif detections and compute z-scores of triadic occurrences analytically (Squartini and Garlaschelli 2011), and an applied study where triadic motifs and their time evolution are used as early warnings of topological collapse during the 2008 financial crisis (Squartini et al. 2013).

Our contribution goes in this direction, using Maximum-Entropy methods constraining degree distributions and strength distributions - in their directed form and taking into account their reciprocal nature - to characterize triadic connections and the total money circulating on them for different product layers of the Dutch production network. An analysis of this kind can give better insight into how much product-level granularity is needed in production network datasets and how the links and weights of a production network are organized for different products. Once product layer patterns have been detected, National Bureau officials - having experience in the domestic trade of that single commodity - can infer if such motifs and anti-motifs are due to commodity-specific characteristics, market imbalances, or represent structures aided by laws. If imbalances and anomalies are detected, they can eventually advance policy laws to nudge a more convenient redistribution of connections and trade volumes.

2 Theoretical Background

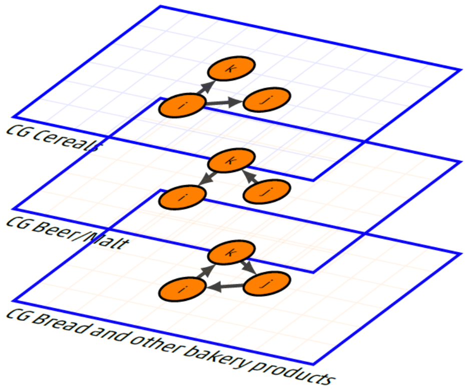


Figure 2.1 Illustrative depiction of a multi-layer network. The nodes i, j and k are connected in different ways through edges, in multiple layers referring to different connection attributes, such as the type of commodity in input/output.This figure is inspired by Di Vece et al. 2023b.

2.1 Networks

A Network $\mathcal{G} \equiv (V, E)$ is a collection of vertices (or nodes) V and edges (or links) E, such that vertices are connected by edges following some predicate which is context-dependent. For example firm i can connect to firm j if they have an intermediate input/output relationship, such as in Production Networks, or they can connect if firm i controls firm j, such as in the case of Control Networks. (Rungi et al. 2017). In both cases a directed connection from i to j is formed. When the network is composed of directed (or oriented) links, we call it *directed*. If, instead, the links are exclusively bi-directional, such as in the case of individuals forming contracts on the Bitcoin Lightning Network (Vallarano et al. 2020), the network is called *undirected*.

The connections are intrinsically characterized by their presence/abscence. A network, where only presence/abscence relationships are present is called *binary*. Instead, if edges are also characterized by a weight, indicating the intensity of the connection, e.g. the amount of output (in monetary value) from firm i to j, the network is called *weighted*.

Weighted Networks are represented in matrix form using a weighted adjacency $N \times N$ matrix $(W)_{ij} = w_{ij}$, where $N = |V|$ is the number of nodes, and w_{ij} is given by the intensity of the connection from node i to node j. It is often useful also to indicated its binary projection $a_{ij} = \Theta(w_{ij})$ reading

$$a_{ij} = \begin{cases} 1, & \text{if node i and j are connected} \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

To describe the structure and the weights of a network and the behavior of its nodes, network statistics are usually computed. These objects are functions of the weighted or/and the binary adjacency matrices.

The degree of a node i is defined as

$$k_i = \sum_{j \neq i} a_{ij} \tag{2}$$

and indicates the number of its connected neighbors. In the case of directed networks the degree is defined for outward links (out-degree) and for inward links (in-degree) appropriately summing on the directed binary adjacency matrix and its transpose. In the case of Production Networks, the out-degree k_i^{out} of firm i counts the number of its users, instead its in-degree k_i^{in} counts the number of its suppliers

$$\begin{cases} k_i^{out} &= \sum_{j\neq i} a_{ij} \\ k_i^{in} &= \sum_{j\neq i} a_{ji}. \end{cases} \tag{3}$$

The analogue of the degree for weighted networks is the strength of node i defined as

$$s_i = \sum_{j\neq i} w_{ij} \tag{4}$$

indicating the sum of the connection intensities of i with its neighbors. In the directed case, the strenght is usually decomposed in out-strength s_i^{out} and in-strength s_i^{in}, expressing the sum of the intensities of connections respectively originating from i and going to i,

$$\begin{cases} s_i^{out} &= \sum_{j\neq i} w_{ij} \\ s_i^{in} &= \sum_{j\neq i} w_{ji}. \end{cases} \tag{5}$$

Finally, networks can be mono-layered, as discussed until now, or multi-layered. In the case of multi-layer networks the connections are labelled by an attribute, specifying the type of connections. For example, consider two industries i and j, connected with each other in different ways in different layers, such as in Fig.2.1: in the layer Cereals, i is a supplier for j, while in the layer 'Bread and other bakery products', it is j that supplies i. A multi-layer approach guarantees additional information with respect to mono-layer aggregates where i and j would be depicted as having a reciprocated connection, exchanging a 'representative' commodity.

2.2 KL Minimization Framework

Depending on our assumption on the Data-Generating Process, we can model the existence of the link and its intensity jointly or separately . In the first case we would want to estimate the distribution of the weighted network W, i.e. $Q(W)$, while in the case of disjoint estimation we want to estimate two objects: the distribution of the binary network A, i.e. $P(A)$, and the distribution of the weighted network W compatible with the binary adjacency matrix A, i.e. $Q(W|A)$, such that $Q(W) = P(A)Q(W|A)$. In the case of our models it is possible to trivially pass from the distribution of weighted networks to the distribution of dyadic weights, and from the distribution of binary networks to the distribution of links.

Separate estimation of links and weights is usually pursued for numerical convenience, given the reduced number of parameters to estimate in a single optimization, and in light of recent results showing a generally good performance of separate (or conditional) methods when compared to joint (or integrated) methods in terms of reproduction of higher-order network statistics. In order to estimate such distributions we can proceed in two different ways

- Proceed by assuming the distribution function and then test the robustness of our results to change in distributional choice, a posteriori;
- Estimating distributions by requiring minimal assumptions.

We follow the second path by maximising the uncertainty of our economic system given some constraints, representing available data. This, in turn, implies a minimization of the Kullback-Leibler Divergence given some constraints, representing the information we have on specific network statistics.

The Kullback-Leibler divergence $D_{KL}(Q||R)$ is a measure of the information lost when we use a posterior distribution Q to approximate a prior distribution R. In mathematical terms it is defined as

$$D_{KL}(Q||R) = \int_W Q(W) \ln\left(\frac{Q(W)}{R(W)}\right) dW \tag{6}$$

where, in our case the integral is on all possible realizations of W, $Q(W)$ is the posterior distribution of weights (and our target distribution) and $R(W)$ is the analoguous prior.

By requiring the factorization assumption on both the posterior $Q(W) = P(A)Q(W|A)$ and the prior $R(W) = T(A)R(W|A)$ we obtain

$$D_{KL}(Q||R) = -S[P] - S[Q|P] + S[P,T] + S[Q,R|P] \tag{7}$$

where

$$S[P] = -\sum_A P(A) \ln(P(A)) \tag{8}$$

the Shannon Entropy related to the binary posterior distribution $P(A)$,

$$S[Q|P] = -\sum_A P(A) \int_{W_A} Q(W|A) \ln(Q(W|A))dW \tag{9}$$

the Shannon Entropy related to the conditional distribution $Q(W|A)$ given P.

The remaining terms, $S[P,T]$ and $S[Q,R|P]$ are cross-entropies of the form

$$S[P,T] = -\sum_A P(A) \ln(T(A)) \tag{10}$$

and

$$S[Q,R|P] = -\sum_A P(A) \int_{W_A} Q(W|A) \ln(R(W|A))dW. \tag{11}$$

which include prior information.

By requiring minimal prior assumptions, i.e. that the priors are uniform with a support equal or greater than the one of the relative posterior, the last two terms are constant and minimizing D_{KL} equals maximising the joint entropy $S[Q]$

$$S[Q] = S[P] + S[Q|P]. \tag{12}$$

The constrained maximisation of the entropy $S[Q]$ leads to the distribution of binary networks $P(A)$ and compatible weighted networks $Q(W|A)$ that on average respect a set of binary constraints $\{C_l(A)\}$ and a set of weighted constraints $\{C_m(W)\}$, where $C_\alpha(A)$ is the l-th constraint, expressed as a function of A, and C_m is the m-th constraint, expressed as a function of W.

For convenience, let us treat the binary problem of constrained maximization, the relative problem for weighted networks can be treated in the same way. Our objective is to estimate $P(A)$ which maximize $S[P]$ given a normalization constraint

$$\sum_A P(A) = 1 \tag{13}$$

and constraints on links $C_l(A)$

$$\sum_A C_l(A)P(A) = C_l^* \tag{14}$$

where C_l^* is the known measure of C_l on the empirical network A^*.

To do so, we define the Lagrangian of the constrained problem, to maximize

$$S[P] - \lambda_0 \left(\sum_A P(A) - 1 \right) - \sum_l \theta_l \left(\sum_A C_l(A)P(A) - C_l^* \right) \tag{15}$$

which leads to

$$P(A) = \frac{e^{-\sum_l \theta_l C_l(A)}}{\sum_A e^{-\sum_l \theta_l C_l(A)}}. \tag{16}$$

In similar way it is possible to show that the solution to the distribution estimation for weights is

$$Q(W|A) = \begin{cases} \dfrac{e^{-\sum_m \beta_m C_m(W)}}{\sum_{W_A} e^{-\sum_m \theta_m C_m(W)}} & \text{if} \quad W \in W_A \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

where $W \in W_A$ indicates that the specific W belongs to the ensemble of weighted networks compatible with A.

3 Methods and Measures

In Eq.16 and Eq.17 we obtained the general form of the distribution of binary adjacency matrices $P(A)$ and the distribution of compatible weighted matrices $Q(W|A)$ that maximize the joint entropy $S[Q]$ (and minimize the KL Divergence given uniform prior).

To explicitly express the related distribution we need to choose the set of constraints $C_l(A)$ and $C_m(W)$ and subsequently find the related set of Lagrange parameters θ_l and β_m using Maximum (Log-)Likelihood Estimation (MLE) for binary models and Generalized Maximum Log-Likelihood Estimation (GLE) for conditional weighted models. For convenience, we leave the mathematical details about the MLE and the GLE estimation of our models in Appendix A and B, respectively.

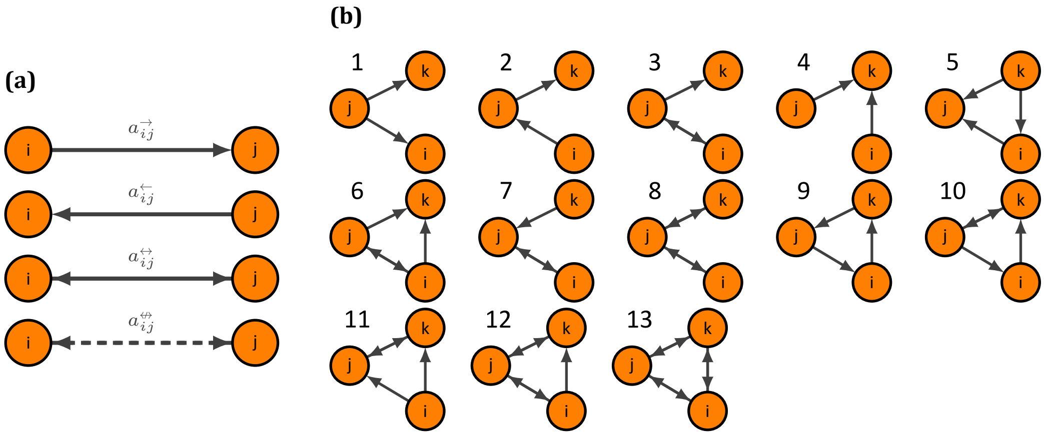


Figure 3.1 (a) Decomposition of generic link (i, j) in non-reciprocated and reciprocated edges. (b) The 13 possible configuration of connected triads, labelled by their m-type. This figure is inspired by the paper in Di Vece et al. 2023b.

3.1 Binary Null Models

In the case of binary models, the constraints that are usually imposed are the out-degree k_i^{out} and in-degree k_i^{in} sequences. We are requiring that in our statistical ensemble of economic networks, both statistics are reproduced in average, allowing for small deviations related to noise or external factors (e.g. interaction with industries in another country). In inter-industry production networks, this implies that we are constraining the number of suppliers and users of a industry i. The implied model is called Directed Binary Configuration Model (DBCM) and is characterized by a probability distribution of the form

$$P(A) = \prod_{i,j\neq i} p_{ij}^{a_{ij}}(1-p_{ij})^{1-a_{ij}} \tag{18}$$

where the dyadic connection probability $Pr(a_{ij} = 1) = p_{ij}$ is

$$p_{ij} = \frac{e^{-\theta_i^{out}-\theta_j^{in}}}{1+e^{-\theta_i^{out}-\theta_j^{in}}}. \tag{19}$$

where θ_i are estimated using MLE on the log-likelihood $\mathcal{L} = \ln(P(A))$, which is equivalent to requiring

$$\begin{cases} k_i^{out;*} & = \sum_{j\neq i} p_{ij} = \langle k_i^{out} \rangle \\ k_i^{in;*} & = \sum_{j\neq i} p_{ji} = \langle k_i^{in} \rangle \end{cases} \tag{20}$$

where $k_i^{out,*}$ is the empirical out-degree of the i-th node and $\langle k_i^{out} \rangle$ is the average over the ensemble of networks of the same measure. The definitions for in-degree sequences are analogous.

If we are interested only in constraining the tendency of industries to have a certain number of suppliers and users, then DBCM would be our preferred model. In the present case, in which we are also interested in reciprocation of links, instead, we have to process our links in a different way: for each couple we differentiate if a given connection is uni-directional or bi-directional, moving from a_{ij} to the set $\{a^{\rightarrow}_{ij}, a^{\leftarrow}_{ij}, a^{\leftrightarrow}_{ij}\}$ as illustrated in Fig.3.1(a). The related degree sequences are of three kinds:

– non-reciprocated out-degree $k_i^{\rightarrow}$, indicating the portion of out-degree of node i which is related to uni-directional links,

$$k_i^{\rightarrow} = \sum_{j \neq i} a_{ij}^{\rightarrow} = \sum_{j \neq i} a_{ij}(1 - a_{ji}) \tag{21}$$

– non-reciprocated in-degree $k_i^{\leftarrow}$, indicating the portion of in-degree of node i which is related to uni-directional links,

$$k_i^{\leftarrow} = \sum_{j \neq i} a_{ij}^{\leftarrow} = \sum_{j \neq i} a_{ji}(1 - a_{ij}) \tag{22}$$

– reciprocated degree $k_i^{\leftrightarrow}$, indicating the portion of out or in-degree of node i which are related to bi-directional links,

$$k_i^{\leftrightarrow} = \sum_{j \neq i} a_{ij}^{\leftrightarrow} = \sum_{j \neq i} a_{ij} a_{ji}. \tag{23}$$

The corresponding model is called Reciprocal Binary Configuration Model (RBCM) and is characterized by the probability distribution

$$P(A) = \prod_{i, j \leq i} (p_{ij}^{\rightarrow})^{a_{ij}^{\rightarrow}} (p_{ij}^{\leftarrow})^{a_{ij}^{\leftarrow}} (p_{ij}^{\leftrightarrow})^{a_{ij}^{\leftrightarrow}} (p_{ij}^{\nleftrightarrow})^{a_{ij}^{\nleftrightarrow}} \tag{24}$$

with connection probabilites with prescriptions

$$\begin{cases} p_{ij}^{\rightarrow} & = \dfrac{x_i^{\rightarrow} x_j^{\leftarrow}}{1 + x_i^{\rightarrow} x_j^{\leftarrow} + x_i^{\leftarrow} x_j^{\rightarrow} + x_i^{\leftrightarrow} x_j^{\leftrightarrow}} \\ p_{ij}^{\leftarrow} & = \dfrac{x_i^{\leftarrow} x_j^{\rightarrow}}{1 + x_i^{\rightarrow} x_j^{\leftarrow} + x_i^{\leftarrow} x_j^{\rightarrow} + x_i^{\leftrightarrow} x_j^{\leftrightarrow}} \\ p_{ij}^{\leftrightarrow} & = \dfrac{x_i^{\leftrightarrow} x_j^{\leftrightarrow}}{1 + x_i^{\rightarrow} x_j^{\leftarrow} + x_i^{\leftarrow} x_j^{\rightarrow} + x_i^{\leftrightarrow} x_j^{\leftrightarrow}} \\ p_{ij}^{\nleftrightarrow} & = \left[1 + x_i^{\rightarrow} x_j^{\leftarrow} + x_i^{\leftarrow} x_j^{\rightarrow} + x_i^{\leftrightarrow} x_j^{\leftrightarrow}\right]^{-1}. \end{cases} \tag{25}$$

where $x_i^{\rightarrow} \equiv e^{-\theta_i^{\rightarrow}}$, $x_i^{\leftarrow} \equiv e^{-\theta_i^{\leftarrow}}$ and $x_i^{\leftrightarrow} \equiv e^{-\theta_i^{\leftrightarrow}}$ are the exponentiated Lagrange multipliers regulating for the non-reciprocated out-degree, non-reciprocated in-degree and reciprocated degree respectively. The parameters are estimated using MLE, maximizing $\mathcal{L} = \ln(P(A))$, which is equivalent to require

$$\begin{cases} k_{i,*}^{\rightarrow} & = \langle k_i^{\rightarrow} \rangle = \sum_{j \neq i} p_{ij}^{\rightarrow} \\ k_{i,*}^{\leftarrow} & = \langle k_i^{\leftarrow} \rangle = \sum_{j \neq i} p_{ij}^{\leftarrow} \\ k_{i,*}^{\leftrightarrow} & = \langle k_i^{\leftrightarrow} \rangle = \sum_{j \neq i} p_{ij}^{\leftrightarrow}, \end{cases} \tag{26}$$

i.e., the equivalence between empirical and ensemble-averaged degrees.

3.2 Conditionally Weighted Null Models

When inquiring the structure of weights. the constraints that are usually imposed are out-strength s_i^{out} and in-strength s_i^{in} sequences. As in the binary case, enforcing these constraints insures that both statistics are reproduced in average in the ensemble of generated networks. Constraining these quantities implies that for each industry i we maintain its output, i.e. the amount given by i to all its users, and its input, i.e. the amount taken by i from all its suppliers, equal to their empirical measures. Such conditional model is called CReM$_A$. The related distribution $Q(W|A)$ reads

$$Q(W|A) = \prod_i \prod_{j \neq i} \left[\left(\beta_i^{out} + \beta_j^{in} \right) e^{-\left(\beta_i^{out} + \beta_j^{in}\right) w_{ij}} \right]^{f_{ij}}, \tag{27}$$

where β_i^{out} and β_i^{in} are Lagrange parameters which are estimated using Generalized Log-Likelihood Estimation (GLE) and $f_{ij} = \langle a_{ij} \rangle$ is the connection probability induced by the binary model of choice. This kind of estimation, also called an 'annealed' procedure (Di Vece et al. 2023a), ensures that binary random variability is taken into account when estimating parameters the distribution of weights. Also GLE requires the solution of a system of first-order equations that equate empirical out-strength and in-strength to ensemble-averaged counterparts:

$$\begin{cases} s_i^{out;*} & = \sum_{j \neq i} w_{ij} = \langle s_i^{out} \rangle \\ s_i^{in;*} & = \sum_{j \neq i} w_{ji} = \langle s_i^{in} \rangle. \end{cases} \tag{28}$$

According to the CReM$_A$ the weight w_{ij} is then estimated as

$$\langle w_{ij} \rangle = p_{ij} \langle w_{ij} | a_{ij} = 1 \rangle \tag{29}$$

where p_{ij} depends on the binary model of choice and the conditional weight $\langle w_{ij} | a_{ij} = 1 \rangle$ is

$$\langle w_{ij} | a_{ij} = 1 \rangle = \left(\beta_i^{out} + \beta_j^{in} \right)^{-1}. \tag{30}$$

We can modify the previous model such that we can take into account the reciprocity profile of links on which the weights are sampled. To do so, we define four types of strengths:

- non-reciprocated out-strength $s_i^{\rightarrow}$, indicating the output of industry i on uni-directional links, namely

$$s_i^{\rightarrow} = \sum_{j \neq i} w_{ij}^{\rightarrow} = \sum_{j \neq i} a_{ij}^{\rightarrow} w_{ij}; \tag{31}$$

- non-reciprocated in-strength $s_i^{\leftarrow}$, indicating the input of industry i on uni-directional links, namely

$$s_i^{\leftarrow} = \sum_{j \neq i} w_{ij}^{\leftarrow} = \sum_{j \neq i} a_{ij}^{\leftarrow} w_{ji}; \tag{32}$$

- reciprocated out-strength $s_i^{out,\leftrightarrow}$, indicating the output of industry i on bi-directional links, namely

$$s_i^{out,\leftrightarrow} = \sum_{j \neq i} w_{ij}^{out,\leftrightarrow} = \sum_{j \neq i} a_{ij}^{\leftrightarrow} w_{ij}; \tag{33}$$

- reciprocated in-strength $s_i^{in,\leftrightarrow}$, indicating the input of industry i on bi-directional links, namely

$$s_i^{in,\leftrightarrow} = \sum_{j \neq i} w_{ij}^{in,\leftrightarrow} = \sum_{j \neq i} a_{ij}^{\leftrightarrow} w_{ji}. \tag{34}$$

Enforcing these network statistics as constraints leads to a separable distribution of weights $(Q(W|A) = Q(W|A^{\rightarrow}) Q(W|A^{\leftrightarrow})$ and, hence, to the resolution of two separate sub-problems: one for the uni-directional links, characterized by

$$Q(W|A^{\rightarrow}) = \prod_{i,j \neq i} \left[\left(\beta_i^{\rightarrow} + \beta_j^{\leftarrow} \right) e^{-\left(\beta_i^{\rightarrow} + \beta_j^{\leftarrow} \right) w_{ij}} \right]^{a_{ij}^{\rightarrow}} \tag{35}$$

and the prescription for the conditional weight given by

$$\begin{cases} \langle w_{ij}^{\rightarrow} | a_{ij}^{\rightarrow} = 1 \rangle & = \left(\beta_i^{\rightarrow} + \beta_j^{\leftarrow}\right)^{-1} \\ \langle w_{ij}^{\leftarrow} | a_{ij}^{\leftarrow} = 1 \rangle & = \left(\beta_i^{\leftarrow} + \beta_j^{\rightarrow}\right)^{-1} \end{cases} \tag{36}$$

and one for the reciprocated links, characterized by the distribution

$$Q(W|A^{\leftrightarrow}) = \prod_{i,j \neq i} \left[\left(\beta_i^{out,\leftrightarrow} + \beta_j^{in,\leftrightarrow}\right) e^{-\left(\beta_i^{out,\leftrightarrow} + \beta_j^{in,\leftrightarrow}\right) w_{ij}} \right]^{a_{ij}^{\leftrightarrow}} \tag{37}$$

and the prescription for the conditional weight given by

$$\begin{cases} \langle w_{ij}^{out,\leftrightarrow} | a_{ij}^{\leftrightarrow} = 1 \rangle & = \left(\beta_i^{out,\leftrightarrow} + \beta_j^{in,\leftrightarrow}\right)^{-1} \\ \langle w_{ij}^{in,\leftrightarrow} | a_{ij}^{\leftrightarrow} = 1 \rangle & = \left(\beta_i^{in,\leftrightarrow} + \beta_j^{out,\leftrightarrow}\right)^{-1}. \end{cases} \tag{38}$$

The set of Lagrange parameters $\{\beta_i^{\rightarrow}, \beta_i^{\leftarrow}, \beta_i^{out,\leftrightarrow}, \beta_i^{in,\leftrightarrow}\}$ are obtained using GLE, leading to the solution of two systems of equations equating ensemble-averaged non-reciprocated strengths and reciprocated strengths to their empirical counterparts. This model takes the name of Conditionally Reciprocal Configuration Model (CRWCM), and is used in conjunction with the RBCM for the inference of weights.

3.3 Pattern Detection Analysis

In Pattern Detection Analysis our goal is comparing patterns in the empirical network to expected patterns in null models constraining a subset of empirical features. If there is no significant deviation from expectation we say that our empirical network is in equilibrium with the statistical ensemble, i.e. the constrained empirical features are the only ingredients needed for the system to have that peculiar pattern. If, instead, the empirical pattern is over or under-represented with respect to expectations, we say that the empirical system is out-of-equilibrium with the statistical ensemble, i.e. the constrained features do not give enough information to encode that peculiar structure.

In our case we will focus on the connected triadic structures, seen as the 13 sub-structures in Fig.3.1(b), given by all the possible ways in which three nodes are connected. We will characterize each sub-structure by two quantities, the abundance N_m, namely the number of subgraph of type m in the graph, and the total flux F_m, namely the total amount of money circulating on subgraphs of type m. For instance, consider the sub-graph of type $m = 6$, its abundance N_6 reads

$$N_6 = \sum_i \sum_{j \neq i} \sum_{k \neq i,j} a_{ij}^{\leftrightarrow} a_{jk}^{\rightarrow} a_{ik}^{\rightarrow}, \tag{39}$$

while its total flux F_6 reads

$$F_6 = \sum_i \sum_{j \neq i} \sum_{k \neq i,j} a_{ij}^{\leftrightarrow} a_{jk}^{\rightarrow} a_{ik}^{\rightarrow} \left(w_{ij} + w_{ji} + w_{jk} + w_{ik}\right). \tag{40}$$

To compare expectations with empirical values we refer to z-score analysis, by defining a z-score measure for the two statistics of interest. The z-score of the abundance N_m of subgraphs of type m, with respect to a generic binary null model is

$$z[N_m] = \frac{N_m - \langle N_m \rangle_{model}}{\sigma[N_m]_{model}} \tag{41}$$

where $\langle N_m \rangle_{model}$ is the model-dependent ensemble-averaged value of the abundance N_m and $\sigma[N_m]_{model}$ is its model-dependent standard error across the ensemble realizations. A similar z-score is defined also for the total flux F_m,

$$z[F_m] = \frac{F_m - \langle F_m \rangle_{model}}{\sigma[F_m]_{model}} \tag{42}$$

where the null model is instead a mixture of binary and conditional weighted models.

In Network Science literature, it is often assumed that a triadic structure is a network 'motif' if the associated z-score is $z > 3$, while it is an 'anti-motif' if the associated z-score is $z < -3$. This convention implicitly assumes that the abundance (and the total flux) of triadic structures approximately follow a normal distribution across the statistical ensemble of graphs. In this work we relax this assumption by estimating the confidence interval of each z-score on 500 sample of the system, generated according to the null model distribution of choice.

4 Results

4.1 Data Pre-Processing

We used the published data about the Dutch production network in 2018, containing 'constructed' relationships among $900\,000$ firms which are labelled by the type of commodity supplied, for 650 commodities. This dataset can be viewed as a multilayer network, where nodes are firms, edges are input/output relationships, and the layers indicate the type of commodity supplier/used, i.e. $a_{ij}^{(\alpha)} = 1$ if firm i supplies commodity α to firm j, and $a_{ij}^{(\alpha)} = 0$ otherwise.

However, it is known that the deterministic procedure used by Statistics Netherlands (SN) in inferring links (and weights) leads to biased network statistics, at least for the network density and degree distributions (Rachkov et al. 2021; Kayzel and Pijpers 2023). For this reason we avoid using the inter-firm network as it is and we proceed taking advantage of the known coherence with the input-output tables at SBI4 industry-classification for 192 commodity-groups.

We do so, by aggregating the firms at the industry SBI5 classification level, and aggregating 650 commodity groups into 192. Obviously, we expect some noise due to the choice of the SBI5 classification instead of the SBI4, however we prefer it to increase the granularity of our sample of industries.

After cleaning for intra-industry trade, which is not salient in the detection of the connected triadic motifs, we obtain a multilayer network of 862 industries and 187 commodity groups.

4.2 Exploratory Analysis

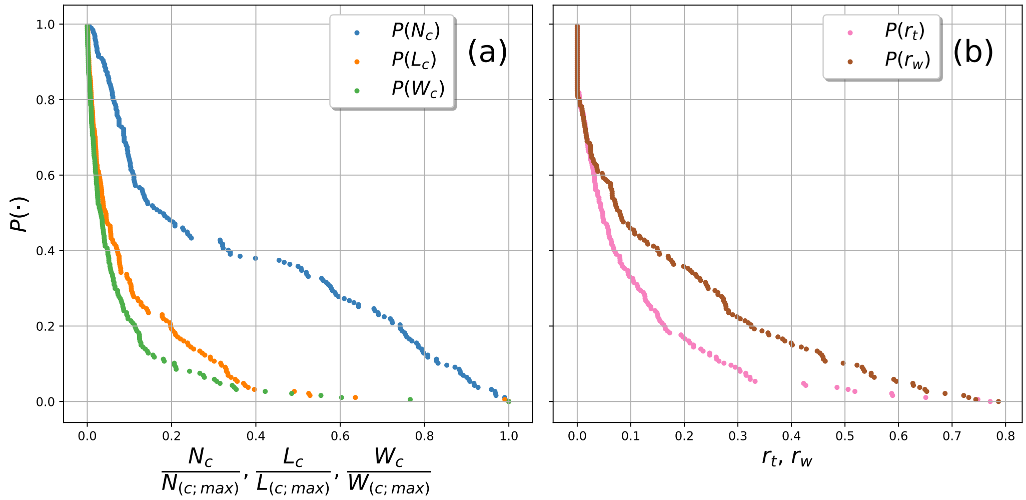


Figure 4.1 Counter-Cumulative distribution for quantities of interest across layers: (a)$P(\cdot)$ for the number of active industries N_c, the connectedness L_c and the total circulating money W_c on commodity c; (b)$P(\cdot)$ for binary reciprocity r_t and weighted reciprocity r_w.

In literature it has been found that in Production Networks firm usually supply(use) to(from) a very specialized set of firms, a notion that would imply that most industries are present in only a few layers, and only a small number of industries are actively engaging in intermediate trade on each layer.

In our case, we can see that this stylized fact is not true, in fact, the counter-cumulative distribution of the number of active industries in a given layer P_N, plotted in Fig.4.1(a), shows a distribution with a very fat tail. Specifically, while it is true that a large number of layers ($> 50\%$) are characterized by a small number of active industries (< 100), it is also the case that a (small) number of commodity groups exist, for which almost the totality of industries is active (~ 818).

In the same figure, we can also investigate the counter-cumulative distribution of the connectedness $L = \sum_{i,j\neq i} a_{ij}$ and total money circulating $W = \sum_{i,j\neq i} w_{ij}$ across layers. Those display a less fat tail with respect to sizes but an uneven distribution across commodities is still present. Specifically, less than 8% of layers have a number of links $L > 6000$, while only 10% of layers have $W > 5000$, i.e. links and volumes, and hence the inter-industry activity in intermediate use/supply, are concentrated on a small amount of commodity groups, with an increased concentration for weights with respect to links.

Let us now move to the analysis of the reciprocal behavior of industries across layers. If industry i and j reciprocate their link, it means that some firms in industry i are suppliers of firms in industry j, while others are their users. The topological reciprocity r_t reading

$$r_t = \frac{L^{\leftrightarrow}}{L} = \frac{\sum_{i,j\neq i} a_{ij}^{\leftrightarrow}}{\sum_{i,j\neq i} a_{ij}} \tag{43}$$

can be seen as a measure of inter-dependence in use/supply between distinct business activities. Its counter-cumulative distribution is depicted in Fig.4.1(b) showing that topological reciprocity is usually very low for most of the commodity groups, while a low number of layers ($< 10\%$) exist for which high reciprocity ($r_t > 0.3$) holds.

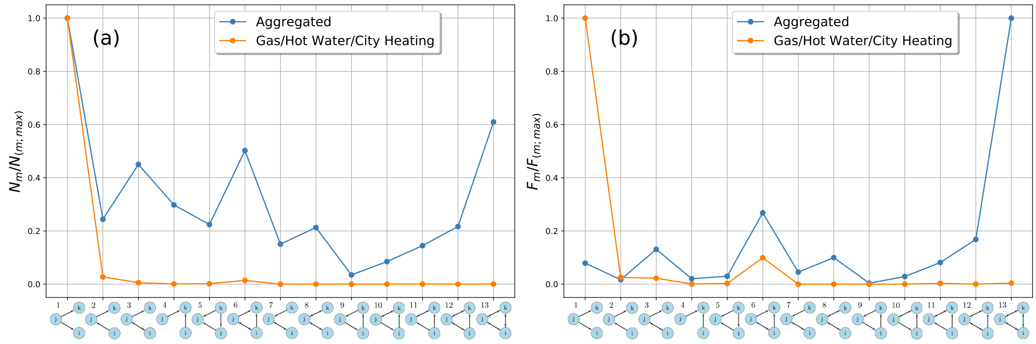


Figure 4.2 Comparison of triadic abundances N_m and triadic fluxes F_m: aggregated network vs 'gas/hot water/city heating commodity. (a) N_m normalized by the maximum across types. (b) F_m normalized by the maximum across types. Slightly modified from Di Vece et al. 2024

A global reciprocity measure can also be defined for the weights, as the weight reciprocity r_w defined as

$$r_w = \frac{W^{\leftrightarrow}}{W} = \frac{\sum_{i,j\neq i} w_{ij}^{\leftrightarrow}}{\sum_{i,j\neq i} w_{ij}} \tag{44}$$

measuring the fraction of weights on reciprocated links.

The counter-cumulative distribution for weighted reciprocity, depicted in Fig.4.1(b) is denoted by a fatter tail than the analogue for links, showing that in a higher amount of commodities (20%) there is a higher weight reciprocity $r_w > 0.3$. This means that while the volume on reciprocated weights is typically low, their intensity is not negligible.

Consider now the abundance N_m and F_m of the 13 triadic subgraphs. In Fig.4.2(a) and Fig.4.2(b) we plot their values for the Aggregated Network (in blue) and the commodity 'Gas/Hot Water/City Heating', to give the reader an example of different behaviours (for triads) arising when the disaggregation at commodity-level is performed. The values are renormalized by the maximum across types for better comparison.

Regarding the intensity of triadic occurrences N_m, we can see in Fig.4.2(a) that while the Aggregated Network has an heterogenous spectrum of triadic types, with the predominance of $m = 1$ (representing open Vs with one suppliers and two users), followed by $m = 13$ (representing fully reciprocated cycles) and $m = 6$, the commodity registers a huge intensity in occurrence of triads of type $m = 1$.

When inspecting the weighted analogue in Fig.4.2(b), for the Aggregated Network the money is highly concentrated around $m = 13$ while in the commodity it is concentrated around $m = 1$.

This specific change does not hold true for all layers: in fact, commodities display a heterogeneous spectrum of commodities, as shown in a very recent paper (Di Vece et al. 2024).

However, the $m = 13$, i.e. the completely reciprocated cycle, is usually an artifact produced by the aggregation of nodes or commodities, that - when a more microscopic description is available - breaks up in favor of more oriented and less reciprocated subgraphs.

From the empirical values N_m and F_m it is not possible to understand if some subgraphs are over, under-represented or possess a lower or higher concentration of money than expected. In the following subsections we will provide some global results on statistical validation of N_m and F_m, while we will more deeply describe only a limited amount of commodity groups of interest, for convenience.

4.3 Pattern Detection Strategy

In the binary case we are interested in measuring the z-score of N_m defined in Eq.41. To do so we need three quantities in total: the empirical N_m, already computed in the previous section, the ensemble-averaged $\langle N_m \rangle$ and its standard deviation $\sigma[N_m]$, according to the null model of choice. In literature (Squartini and Garlaschelli 2011) it is shown that an analytical approach exists to provide exact estimations of the z-score for the null models DBCM and RBCM. While intriguing, the use of the (-3,3) as the confidence interval for validation, implicitly require that N_m is distributed as a normal across ensemble binary configurations, given m. We, instead, find that sampling 500 configurations those distributions are not normal according to a Shapiro Test with 5% confidence intervals. This leads us to use a sampling method to extract the averages and the standard deviation of N_m. To do so, we extract the 2.5 and 97.5 percentiles of N_m out of distribution of 500 realizations. The induced confidence interval will be given by $(z^-[N_m], z^+[N_m])$ where

$$\begin{cases} z^-[N_m] &= \dfrac{N_m^{(2.5)} - \langle N_m \rangle}{\sigma[N_m]} \\ z^+[N_m] &= \dfrac{N_m^{(97.5)} - \langle N_m \rangle}{\sigma[N_m]} \end{cases}. \tag{45}$$

A subgraph of type m is over-represented, and hence a network motif, if $z[N_m] > z^+[N_m]$ while it is under-represented, and hence a network anti-motif, if $z[N_m] < z^-[N_m]$.

In the weighted case we are interested in measuring the z-score of F_m defined in Eq.42. To do so we need three quantities in total: the empirical F_m, already computed in the previous section, the ensemble-averaged $\langle F_m \rangle$ and its standard deviation $\sigma[F_m]$, according to the null model of choice. Also F_m are not normally distributed across the ensemble, hence, also in this case we proceed numerically. We extract the statistics $\langle F_m \rangle$ and $\sigma[F_m]$ from 500 realization of the weighted matrices according to the mixture model DBCM+CReMa, filtering for directed degree and strength sequences, and according to the mixture model RBCM+CRWCM, filtering for degree and strength sequences in their reciprocated and non-reciprocated fashion. The induced confidence interval for $z[F_m]$ will be given by $(z^-[F_m], z^+[F_m])$ where

$$\begin{cases} z^-[F_m] &= \dfrac{F_m^{(2.5)} - \langle F_m \rangle}{\sigma[F_m]} \\ z^+[F_m] &= \dfrac{F_m^{(97.5)} - \langle F_m \rangle}{\sigma[F_m]} \end{cases}. \tag{46}$$

A subgraph of type m has an 'unexpectedly high' total amount of circulating money if, and hence it is a network motif, if $z[F_m] > z^+[F_m]$ while it has a 'unexpectedly low' amount of circulating money, and hence it is a network anti-motif, if $z[F_m] < z^-[F_m]$.

In the following we will use directly the reciprocal null model RBCM+CRWCM given its better performance in identifying binary and weighted triads with respect to DBCM+CReM$_A$ as shown in Di Vece et al. 2024.

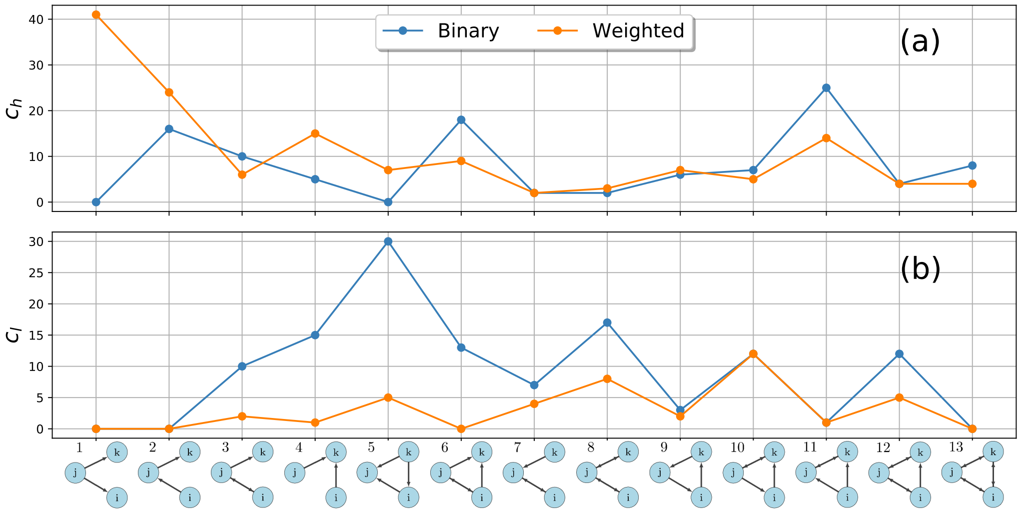


Figure 4.3 Number of commodities having a certain type of (a) motif c_h and (b) anti-motif c_l: comparison of behaviours between binary (in blue) and weighted triadic deviations (in orange). This image is a modification of one in Di Vece et al. 2024

In Fig.4.3(a) the number of commodities c_h having a m-type triadic binary (in blue) and weighted (in orange) motif is depicted. While binary and weighted motif are present with the same type in a restricted number of commodities, $m = 1$ shows a characteristic behaviour. In fact, the formation with one supplier and two users is statistically expected in occurrence, i.e. its amount is due to node-specific and reciprocal tendencies of users and suppliers, however those triads accommodate an unexpectedly high amount of money. For 40 commodities out of 187, users take a large part of their input from selected suppliers, leading to possible systemic propagations if the supply volume of such suppliers is reduced due to exogeneous or endogenous shocks.

On the other hand in Fig.4.3(b) we can see the number of commodities c_l having a m-type of triadic binary (in blue) and weighted (in orange) anti-motif. A high number of commodities have binary anti-motifs of type $m = 5$, followed by $m = 8$. However , while the number of binary anti-motifs is substantial, weighted anti-motifs are reduced in number, i.e. structures can under-occur but the amount of money on them is usually coherent with expectations.

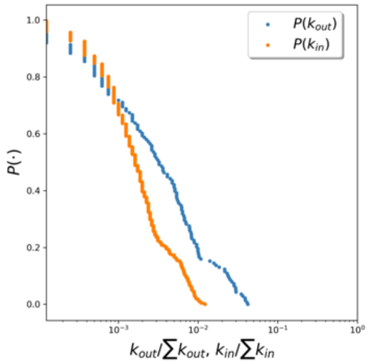
4.4 Triads in Strategic Commodities

In the following we analyze the 'fingerprint' of triadic structures for strategic commodities in the Netherlands.

The first three commodities considered contain firms in the semiconductor sector, which has experienced high demand and global reach since the internet revolution. For food-related commodities, we consider cocoa/chocolate, which has non-trivial structures, as previously identified by CBS. For logistics, we consider "Other Road Transport & Pipelines," which contains a large number of firms that do not specialize in a narrow range of transportation services related to trams, buses, or taxis. For the law sector, we consider "Tax Consultancy," which is highly resonant with the practice of optimising taxation payments and is therefore crucial in fiscal policy. Finally, for utilities, we study the ”Electricity” commodity, which has become increasingly strategic in recent crises, such as the EuroDebt crisis of 2012, the COVID-19 pandemic, and the current Russia-Ukraine war.

4.4.1 Other Machine Devices/Components

The first commodity layer that we enquire is ’Other Machine Devices/Components’. The aforementioned layer consists of users and suppliers of machine devices and components that are not exclusively Metal Components for Doors & Windows, Electrical components, Computer or Peripherals, Telephones and Telecommunication systems, Medical instruments, photo devices or eletrical machinery. In this section, then, falls the production of products such as semiconductors. Given the large impact of the semiconductor industry in the Netherlands it is of interest to describe the nature of connections of the related micro-industries. This commodity layer counts 138 suppliers annd 388 users.



Top 5 industries for out-degree, in-degree, supply and use activity.

Top 5 Out-Degree		Top 5 In-Degree	
Description	Activity	Description	Activity
Man. of other machines/tools for specific purposes	4.2%	Repair/maintenance of general purpose machines and machine parts (not tools)	1.2%
Man. of other machines or equipment for general purposes	4.1%	Repair/maintenance of machines for a specific industry	1.2%
Man. of machinery/equipment for agriculture and forestry	3.9%	General civil and non-residential construction	1.1%
Man. of hoisting, lifting and transport equipment	3.9%	Installation of general purpose machines (no tools)	1.1%
Man. of machines/equipment for industrial refrigeration and air conditioning	3.6%	General metal working	1%

Top 5 Supply		Top 5 Use	
Description	Activity	Description	Activity
Man. of other machines/tools for specific purposes	14%	Man. of other machines/tools for specific purposes	18%
Man. of other machines/tools for general use	13%	General civil and non-residential construction	9%
Man. of hoisting, lifting and transport equipment	10%	Man. of hoisting, lifting and transport equipment	7%
Man. of machines/tools for agriculture and forestry	8%	Man. of other machines/tools for general purposes	6%
Man. of machines for the production of food and beverages	7%	Man. of machines/tools for agriculture and forestry	5%

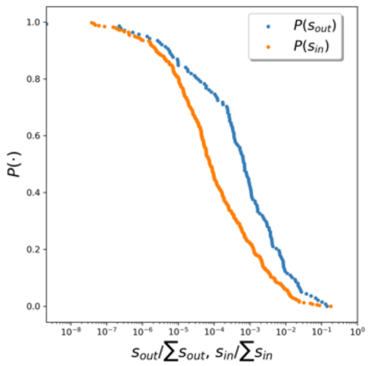


Figure 4.4 (Top-Left) Counter-cumulative distribution of out-degrees and in-degrees for the different industries in the layer ’Other Machine Devices/Components’, with (Top-Right) the top 5 ranked industries in terms of binary activity ($k_{i;out}/\sum_i k_{i;out}$ or $k_{i;in}/\sum_i k_{i;in}$). (Bottom-Left) Counter-cumulative distribution of supply and use volumes for the different industries in the layer ’Other Machine Devices/Components’, with (Bottom-Right) the top 5 ranked industries in terms of supply/use activity ($s_{i;out}/\sum_i s_{i;out}$ or $s_{i;in}/\sum_i s_{i;in}$). Out-going links are more concentrated on a limited number of industries with respect to in-going links. The same is true in average also for supply and use, but there is an industry, i.e. ’Man. of other machine/tools for specific purposes’ that serve as a hub in use activity.

In Fig.4.4 the counter-cumulative distribution for out-degrees and in-degrees (on the top-left panel) and supply and use (bottom-left panel) are depicted. For ’out’ statistics the distribution is plotted in blue, while the analogue for ’in’ statistics are plotted in orange. In both cases, the statistics of interest is normalized by the sum across industries in the commodity layer.

Given that sums are equivalent for out-going and in-going links, i.e. $\sum_i k_{i;in} = \sum_i k_{i;out}$ and $\sum_i s_{i;in} = \sum_i s_{i;out}$ we can compare the distributions. The in-degree distribution reveals a large presence of industries with very few connnections to suppliers with respect to the connections with users. This result is also visible on the right panel for the level of activity ($k_{i;out}/\sum_i k_{i;out}$ and $k_{i;in}/\sum_i k_{i;in}$), where the identity of the top 5 ranked industries is revealed: while the activity of the top 5 ranked industries for out-degree ranges from 3.6% to 4.2%, the same for the in-degree statistics range from 1% to 1.2%.

The distributions for out-strength and in-strength reveal a different behaviour. There is a relative larger number of industries with low use volume with respect to supply volume, but there is also the presence of an industry, specifically 'Manifacture of other machines/tools for specific purposes', which represents a 'use-hub'. In fact, in the bottom-right panel, we can see that its activity is 18%, well above the range of activity in supply volumes, ranging from 7% to 14%.

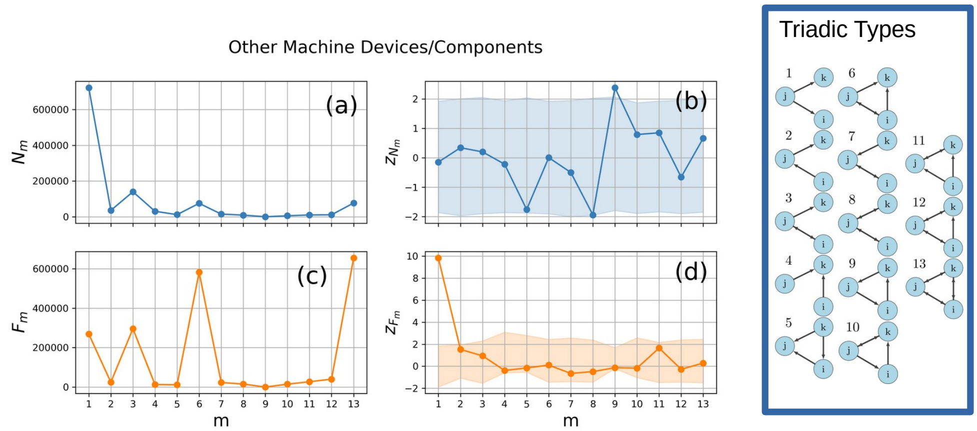


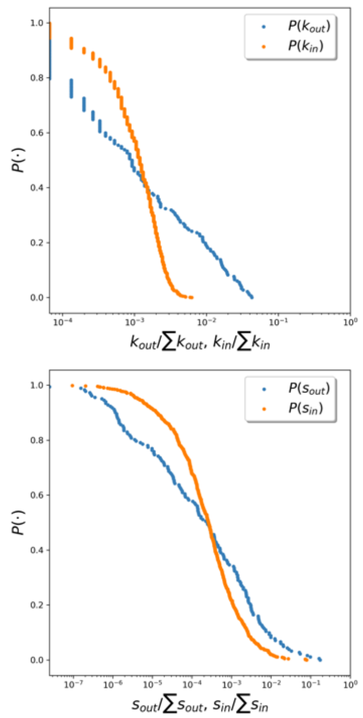
Figure 4.5 (a) N_m and (b) binary triadic pattern detection for the commodity 'Other Machine Devices/Components'. (c) F_m and (b) weighted triadic pattern detection for the same commodity. In occurrence terms $m = 1$, i.e. open-triads with one supplier and two users, are predominant, while being expected according to the RBCM, discounting node-specific and reciprocal tendencies. The binary pattern detection reveals a motif in occurrence for $m = 9$, representig forward feedback loops. Instead, F_m clarifies that on $m = 6$ and especially on $m = 13$ lie a large amount of money. However both concentrations are statistically identified by the reciprocal null model, which actually signals a motif for $m = 1$.

After having explored the activity in supply and use in their binary and weighted forms, we can delve into the triadic pattern detection analysis. In Fig.4.5(a) and (b) we plot the empirical values of N_m and the z-score profile obtained comparing the empirical model to the RBCM. The empirical analysis unveils a high occurrence of subgraph $m = 1$, representing triadic formations connecting a supplier to two users (the different formations are depicted on the right panel). However these triads are within expectation if node-specific properties are discounted. Instead a motif arises for $m = 9$ representing feedback forward loops. A higher than expected presence of these loops have been connected to systemic risk in financial networks, where they represent landing chains that collapse as soon one of the actors suffers a negative shock Squartini et al. 2013. Moving to the weighted triadic flux detection in Fig.4.5(c) and (d), we show the empirical value of F_m across the 13 subgraphs and the related z-score z_{F_m} obtained comparing the empirical network with the reciprocal model RBCM+CRWCM. The empirical investigations reveals a high concentration of money on triads of type $m = 6$ and $m = 13$, however both amounts are within expectations according to the reciprocal model. The z-score profile reveals a higher than expected amount of money on $m = 1$ subgraphs.

Since this is an analysis at the industry-level, feedback forward loops can be thought of as an impending risk due to 'bankruptcy', if and only if the industries interacting in such loops contain a very low number of firms, such that an eventual disconnection of the industry as a whole becomes realistic. Conversely, the weighted for $m = 1$ unveils that a limited number of suppliers retain control of the market, due to properties that cannot be discounted by node-specific or reciprocal factors.

4.4.2 Repair/Installation/Maintenance

The second commodity layer selected for testing is the "Repair/Installation/Maintenance" layer, which consists of suppliers and users of this service. The commodity layer consists of 182 suppliers and 738 users.



Top 5 industries for out-degree, in-degree, supply and use activity.

Top 5 Out-Degree		Top 5 In-Degree	
Description	Activity	Description	Activity
Repair/maintenance of machines for a specific industry	4.3%	General government administration	0.6%
Repair/maintenance of general purpose machines and machine parts (not tools)	4.2%	General civil and non-residential construction	0.6%
Repair/maintenance of ships	4%	Forwarders, shipbrokers, charterers and other intermediaries	0.6%
Installation of electrical equipment	3.8%	Services for arable and/or horticulture	0.5%
Installation of general purpose machines	3.7%	Inland shipping	0.5%

Top 5 Supply		Top 5 Use	
Description	Activity	Description	Activity
Repair/Maintenance of machines for a specific industry	17%	Construction of sports and recreational vessels	8%
Repair/Maintenance of general purpose machines	16%	General government administration	7%
Repair/Maintenance of ships	10%	Passenger air transport	7%
Installation of electrical equipment	8%	Holdings (not financial)	3%
Man. of hoisting, lifting and transport equipment	5%	Manufacture of iron and steel and ferroalloys	2%

Figure 4.6 (Top-Left) Counter-cumulative distribution of out-degrees and in-degrees for the different industries in the layer 'Repair/Installation/Maintenance', with (Top-Right) the top 5 ranked industries in terms of binary activity ($k_{i;out}/\sum_i k_{i;out}$ or $k_{i;in}/\sum_i k_{i;in}$). (Bottom-Left) Counter-cumulative distribution of supply and use volumes for the different industries in the layer 'Repair/Installation/Maintenance', with (Bottom-Right) the top 5 ranked industries in terms of supply/use activity ($s_{i;out}/\sum_i s_{i;out}$ or $s_{i;in}/\sum_i s_{i;in}$). The top 5 ranked industries in-degree activity are not the same as the ones having larger use volumes, with the exception of the 'General Government administration'. Out-degree activity is larger than in-degree activity, meaning that a small number of suppliers connect to a large number of users, with the top supplier being 'Repair/Maintenance of machines for a specific industry' both in binary annd weighted activity.

In Fig.4.6 the counter-cumulative distribution for out-degrees and in-degrees (on the top-left panel) and supply and use (bottom-left panel) are depicted. For 'out' statistics the distribution is plotted in blue, while the analogue for 'in' statistics are plotted in orange. In both cases, the statistics of interest is normalized by the sum across industries in the commodity layer.

Also in this case the in-degree distribution reveals a large presence of industries with very few connections to suppliers with respect to connections with users. This effect is even more pronounced here than in the previous commodity, in fact on the right panel, we can see that the activity in in-degree, i.e. $k_{i;in}/\sum_i k_{i;in}$, for the top 5 industries ranges from 0.5% to 0.6%. The user industries with more suppliers are the those pertaining to the 'General Government Administration' and 'General civil and non-residential construction'.

Interestingly enough, the industries with more suppliers are not the ones with larger use volumes: on the bottom right panel, in fact, we can see that the top 5 ranked industries for use volumes is different with the exception of the 'General Government administration'. The top ranked industry is 'Construction of sports and recreational vessels', which has a small number of suppliers but the highest amount of use volume.

On the supply volume side we can identify a higher concentration of activity in the top 5 ranked industries, ranging from 5% to 17%, with the most important supplier being 'Repair/Maintenance of machines for a specific industry'.

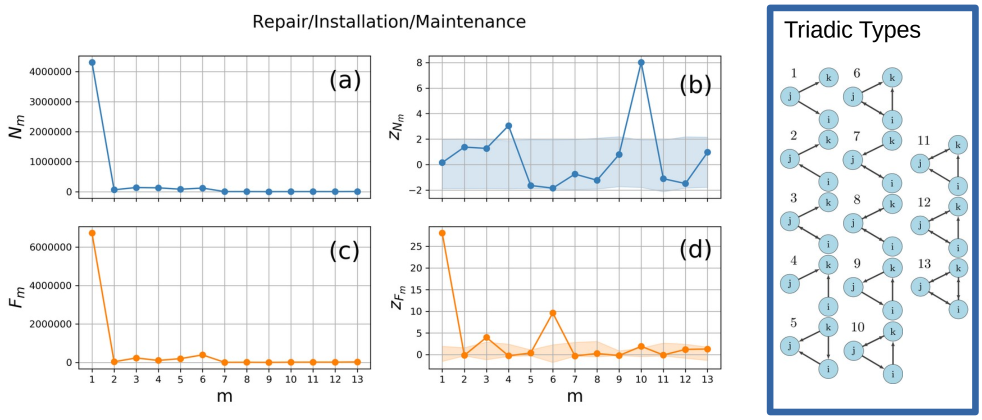


Figure 4.7 (a) N_m and (b) binary triadic pattern detection for the commodity 'Repair/Installation/Maintenance'. (c) F_m and (b) weighted triadic pattern detection for the same commodity. The triad-type $m = 1$, i.e. open-triads with one supplier and two users, is predominant both in occurrence and flux terms. In absolute terms $m = 1$ has a high number of occurrences with a large amount of money. Triadic pattern detection signals $m = 4$ and $m = 10$ as binary motifs while it signals $m = 1$ and $m = 6$ as weighted motifs.

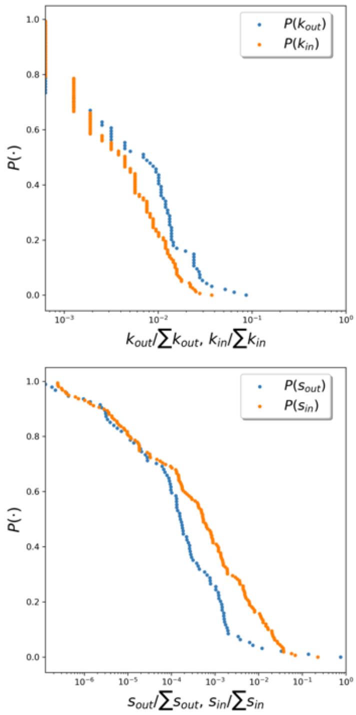
Moving to the triadic pattern detection analysis, in Fig.4.7(a) the N_m and (b) the related z-scores z_{N_m} are depicted. The study of occurrences signal a large presence of subgraph $m = 1$, however this behaviour is within expectations when node-specific properties are discounted. Instead, the z-score profile signals to motifs for $m = 4$, i.e. the open triad connecting two suppliers to the same user, and $m = 10$, i.e. a cyclic closed loop with a bi-directional link. The larger presence of $m = 4$ can be thought of the commodity-specific 'slight' vulnerability to demand shock, in case the firms within the user industry go bankrupt and are hence disconnected from the market.

When inspecting weights on triads through F_m, in Fig.4.7(c) we see that a large amount of money rest on formations of type $m = 1$, moreover this amount is unexpectedly high according to the null model, as extracted from the z-score profile in Fig.4.7(d).

The weight analysis gives us the information that while the layer is vulnerable to demand shocks in case of bankruptcy it is also vulnerable to supply shock in volumes, i.e. a reduction in supply volumes of suppliers in formations $m = 1$. Furthermore, the weighted z-score profile also signals a significant motif form $m = 6$, a formation with two suppliers/users and one exclusive user.

4.4.3 Metal Components for Doors & Windows

The third commodity layer selected for testing is the ”Metal Components for Doors & Windows” layer, which consists of suppliers and users of this product. This commodity was chosen because of its unusual weighted triadic structure and the presence of highly important firms. The aforementioned layer consists of 94 suppliers and 159 users.



Top 5 industries for out-degree, in-degree, supply and use activity.

Top 5 Out-Degree		Top 5 In-Degree	
Description	Activity	Description	Activity
Man. of metal structures and parts thereof	8.6%	General civil and non-residential construction	3.7%
Man. of metal doors, windows and frames	6.4%	Man. of metal structures and parts thereof	2.8%
Engineers and other technical design	5.2%	Other specialized work in construction	2.5%
General metalworking	3.7%	Electrical construction installation	2.5%
Man. of machines and equipment for industrial refrigeration and air conditioning	3.3%	Plumbing and fitting, installation of sanitary facilities	2.3%

Top 5 Supply		Top 5 Use	
Description	Activity	Description	Activity
Man. of metal structures and parts thereof	76%	General civil and non-residential construction	23%
Man. of metal doors, windows and frames	14%	Road Construction	7%
Engineers and other technical design	3.3%	Man. of other machines/tools for specific purposes	5.7%
General metalworking	1.2%	Other specialized work in construction	3.7%
Man. of steel tubes/pipes	0.7%	General government administration	3.7%

Figure 4.8 (Top-Left) Counter-cumulative distribution of out-degrees and in-degrees for the different industries in the layer ’Metal Components for Doors & Windows’, with (Top-Right) the top 5 ranked industries in terms of binary activity ($k_{i;out}/\sum_i k_{i;out}$ or $k_{i;in}/\sum_i k_{i;in}$). (Bottom-Left) Counter-cumulative distribution of supply and use volumes for the different industries in the layer ’Metal Components for Doors & Windows’, with (Bottom-Right) the top 5 ranked industries in terms of supply/use activity ($s_{i;out}/\sum_i s_{i;out}$ or $s_{i;in}/\sum_i s_{i;in}$). Supply is monopolized both in binary and weighted activity by the industry ”Man. of metal structures ad parts thereof”, counting for the 8.6% of out-degree activity and 76% of supply activity, while on the demand side ”General civil and non-residential construction” is a quasi-monopoly counting for 3.7% of in-degree activity and 23% of use activity.

In Fig.4.8 the counter-cumulative distribution for out-degrees and in-degrees (on the top-left panel) and supply and use (bottom-left panel) are depicted. For ’out’ statistics the distribution is plotted in blue, while the analogue for ’in’ statistics are plotted in orange. In both cases, the statistics of interest is normalized by the sum across industries in the commodity layer.

The out-degree and in-degree distributions (top-left panel) provide a qualitatively similar behaviour to the previous commodities: the out-degree distribution provide a fatter tail than the in-degree distribution, i.e. a larger number of industries have few ingoing connections (less suppliers) with respect to outgoing connections (more users). In contrast, the in-strength and outgoing distributions have the inverse behaviour: a higher number of industries accommodate a higher amount of use volumes. This holds true with the exception of the top ranked industry in out-degree activity and supply volume, i.e. 'Manifacturing of metal structures annd parts thereof' which concentrates the 8.6% of out-going connections and the 76% of total supply. A similar, but less quantitatively important, concentration is revealed also on the demand side with the industry "General Civil annd non-residential construction" which accomodates the 3.7% of in-going links and the 23% of use volumes.

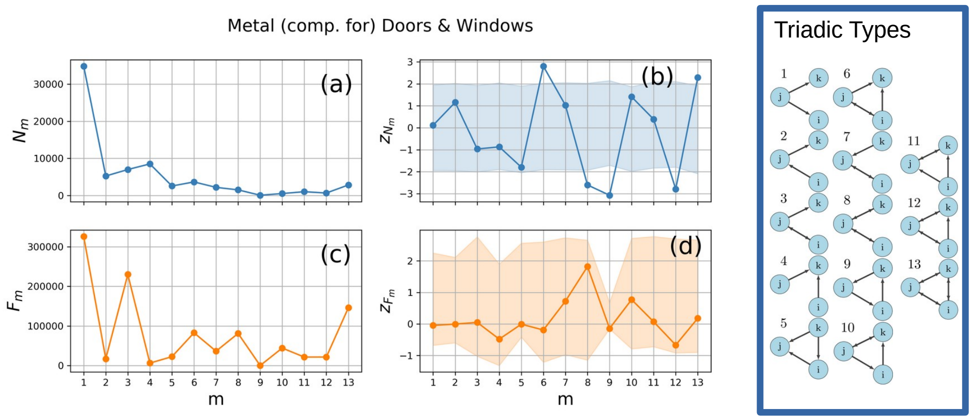


Figure 4.9 (a) N_m and (b) binary triadic pattern detection for the commodity 'Metal Components for Doors & Windows'. (c) F_m and (d) weighted triadic pattern detection for the same commodity. In absolute values $m = 1$ accomodates the higher occurrence and concentration of money. However, these findings are within expectations according to the reciprocal null model. Binary pattern detection signals a large number of motifs ($m = 6$ and $m = 13$) and anti-motifs ($m = 8$, $m = 9$ ad $m = 12$). However, even if occurrences for different triads are unexpected, the amount of money on them is within expectation.

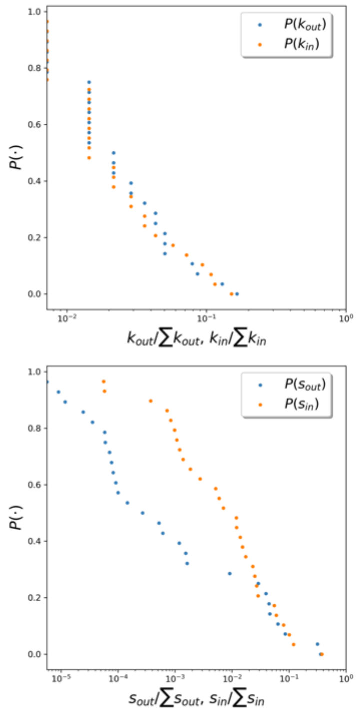
Figure 4.9 depicts the (a) number of triadic occurrences, N_m, and the (c) amount of triadic fluxes, F_m, on connected triples. In terms of occurrence, the predominant structures are $m = 1$ (predominant) and $m = 4$. These types are associated with open triads with one supplier and two users or one user and two suppliers, which leaves the commodity vulnerable to both user and supply shocks in the event of bankruptcy. In weighted terms, there is a high concentration of money on many types, namely $m = 1$, $m = 3$, and $m = 13$, without a single type predominating. It is noteworthy that there is a relatively high concentration of money in totally reciprocal cycles, i.e. where suppliers are also users and supply and use cyclically in triadic formation in a bi-directional manner.

When comparing the empirical system with the null models in Fig.4.9(b), we find that only binary motifs arise. Specifically, there are two motif ($m = 6$ and $m = 13$) and three anti-motifs ($m = 8$, $m = 9$, and $m = 12$). However, both negative and positive z-score are very close to the interval $(-2, 2)$ at the 5% significance level.

At the same time no weighted motifs or anti-motifs are present in Fig.4.9(d), i.e. the statistical null model is able to recover the amount of money on different triads from node-specific properties.

4.4.4 Cocoa/Chocolate

The fourth commodity layer that we choose for our testing is the 'Cocoa/Chocolate' layer, consisting of suppliers and users of the aforementioned product, as representative of Food commodities. This layer consists of 28 suppliers and 29 users, making it a very small layer for number of industries involved.



Top 5 industries for out-degree, in-degree, supply and use activity.

Top 5 Out-Degree		Top 5 In-Degree	
Description	Activity	Description	Activity
Man. of chocolate and confectionery	16.5%	Man. of chocolate and confectionery	15.1%
Processing of cocoa	12.9%	Cafés	11.5%
Man. of bread and fresh pastry	8.6%	Restaurants	10.8%
Man. of spices, sauces and seasonings, salt refining	7.9%	Fast food restaurants	9.3%
Processing of fruit and vegetables	5%	Event catering	7.2%
Top 5 Supply		**Top 5 Use**	
Description	Activity	Description	Activity
Man. of chocolate and confectionery	36%	Man. of chocolate and confectionery	37.9%
Processing of cocoa	31.7%	Man. of long-life pastry products	12%
Man. of other foodstuffs	8.6%	Man. of bread and fresh pastry	10%
Man. of spices, sauces, seasonings, salt refining	6.4%	Man. of dairy products	8.1%
Processing of coffee and tea	4.6%	Restaurants	6%

Figure 4.10 (Top-Left) Counter-cumulative distribution of out-degrees and in-degrees for the different industries in the layer 'Cocoa/Chocolate', with (Top-Right) the top 5 ranked industries in terms of binary activity ($k_{i;out}/\sum_i k_{i;out}$ or $k_{i;in}/\sum_i k_{i;in}$). (Bottom-Left) Counter-cumulative distribution of supply and use volumes for the different industries in the layer 'Cocoa/Chocolate', with (Bottom-Right) the top 5 ranked industries in terms of supply/use activity ($s_{i;out}/\sum_i s_{i;out}$ or $s_{i;in}/\sum_i s_{i;in}$). In all cases, both binary and weighted and both supply and use, there is a monopolizing industry, i.e. 'Man. of chocolate and confectionary'.

In Fig.4.10 the counter-cumulative distribution for out-degrees and in-degrees (on the top-left panel) and supply and use (bottom-left panel) are depicted. For 'out' statistics the distribution is plotted in blue, while the analogue for 'in' statistics are plotted in orange. In both cases, the statistics of interest is normalized by the sum across industries in the commodity layer.

Given the low number of industries in this commodity layer, the distributions contain only a small number of points. The out-degree and in-degree distributions have a similar behaviour, implying a similar concentration of in-degree and out-degree activities across industries. Taking the top 5 ranked industries (for activity) we see that they accommodate a number of suppliers ranging from 7.2% to 15.1% and a number of users ranging from 5% to 16.5% of the total. In both cases the monopolizing industry is 'Manifacturing of chocolate and confectionary'. The same industry monopolizes also supply and use volumes with a 36% of supply activity and a 37.9% of use activity.

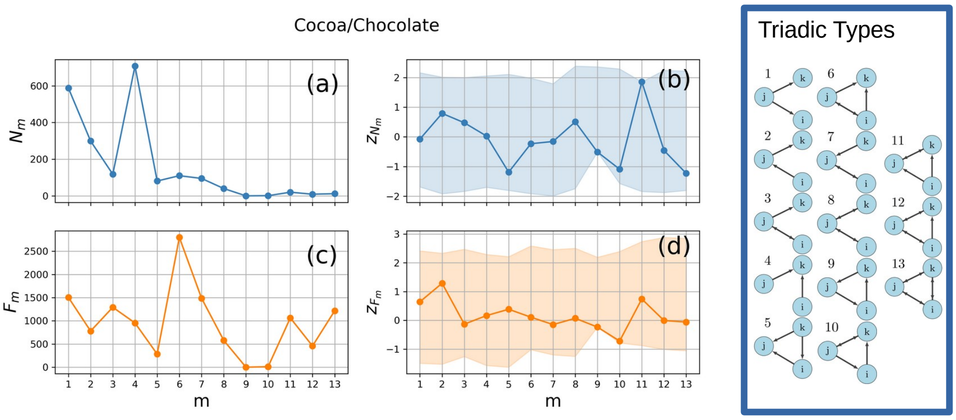


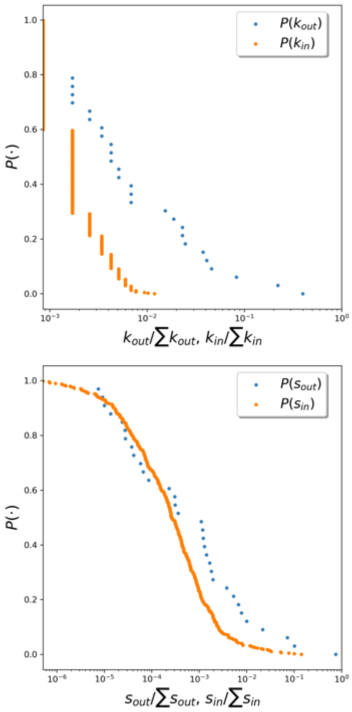
Figure 4.11 (a) N_m and (b) binary triadic pattern detection for the commodity 'Cocoa/Chocolate'. (c) F_m and (b) weighted triadic pattern detection for the same commodity. While triads of type $m = 1$ and $m = 4$ occur with a high amount, their occurrence is within expectations according to the null model. The same happens for $m = 6$ in terms of money concentration on triads.

In Fig.4.11(a) the number of triadic occurrences N_m and (c) amount of triadic fluxes F_m on the connected triples are depicted. In terms of occurrence the predominant structures are $m = 1$ and $m = 4$, the latter formed by two suppliers connected to the same user. Instead, in weighted terms, the largest concentration of money lies on $m = 6$, formed by a bi-directional connection between two industries that are both suppliers of one user. The presence of a bi-directional connection signals the presence of suppliers that are also users in the same commodity layers. A possibility that is discarded when approximating commodity layers with business sectors, as usually done in abscence of survey data.

When we compare the empirical system with the null models in Fig.4.11(b) and (d), we show that no motif arise for triadic occurrences and triadic fluxes. The heterogeneous profile of N_m and F_m is hence well predicted by both RBCM and RBCM+CRWCM, i.e. input/output tendencies of single micro-industries on reciprocated and non-reciprocated links totally describe the mentioned statistics.

4.4.5 Other Road Transport & Pipelines

The commodity layer 'Other Road Transport & Pipelines' comprises micro-industries that supply or use transport services that are not exclusively related to Tram, Bus or Taxi services. This commodity layer consists of 33 suppliers and 476 users.



Top 5 industries for out-degree, in-degree, supply and use activity.

Top 5 Out-Degree		Top 5 In-Degree	
Description	Activity	Description	Activity
Freight transport by road	39.9%	Freight transport by road	1.2%
Moving transport	22.1%	Supermarkets and similar stores with a general range of food and beverages	1%
Services for arable and/or horticulture	8.2%	Land transport services	0.9%
Services for breeding and keeping animals	4.6%	Shops in kitchens	0.8%
Travel organization	4.1%	Forwarders, shipbrokers, charterers and other intermediaries in freight transport	0.8%

Top 5 Supply		Top 5 Use	
Description	Activity	Description	Activity
Freight transport by road	75.7%	Freight transport by road	14.3%
Moving transport	10.2%	Collection of harmless waste	10%
Services for arable and/or horticulture	7.2%	Treatment of harmless waste	10%
Transport via pipelines	2.2%	Petroleum	5.4%
Intermediaries in freight transport	1%	Moving transport	5.2%

Figure 4.12 (Top-Left) Counter-cumulative distribution of out-degrees and in-degrees for the different industries in the layer 'Other Road Transport & Pipelines', with (Top-Right) the top 5 ranked industries in terms of binary activity ($k_{i;out}/\sum_i k_{i;out}$ or $k_{i;in}/\sum_i k_{i;in}$). (Bottom-Left) Counter-cumulative distribution of supply and use volumes for the different industries in the layer 'Other Road Transport & Pipelines', with (Bottom-Right) the top 5 ranked industries in terms of supply/use activity ($s_{i;out}/\sum_i s_{i;out}$ or $s_{i;in}/\sum_i s_{i;in}$).

In Fig.4.12 the counter-cumulative distribution for out-degrees and in-degrees (on the top-left panel) and supply and use (bottom-left panel) are depicted. For 'out' statistics the distribution is plotted in blue, while the analogue for 'in' statistics are plotted in orange. In both cases, the statistics of interest is normalized by the sum across industries in the commodity layer.

A very limited number of suppliers are connected to a plethora of users. On the top-left panel we can see a large number of users with a relatively very small number of connections. Conversely, the out-degree distribution has a fatter tail. On the bottom-left panel the out-strength and in-strength distributions are very similar for low values of the statistics, while they differ for high values, where a large number of industries have higher supply volumes.

In both use and supply (both binary and weighted) the industry taking the largest activity is 'Freight tranport by road', taking into account all firms that transport exclusively on the road. The number of its users is the 39% of the total, with the 75% in supply volume activity. Conversely, its number of suppliers is the 1.2% of the total, with the 14% for use volume activity. In terms of supply its activity constitutes a quasi-monopoly where the second ranked is the industry 'Moving Transport' with the 22.1%(compared to 39%) in out-degree activity and a mere 10.2%(compared to 75%) in supply volume activity.

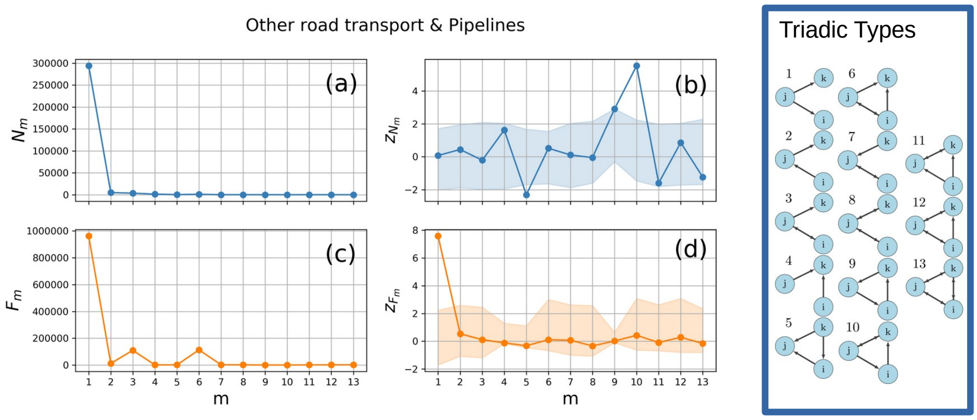


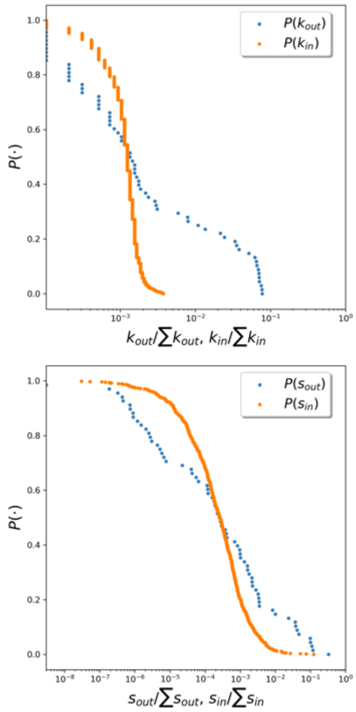
Figure 4.13 (a) N_m and (b) binary triadic pattern detection for the commodity 'Other Road Transport & Pipelines'. (c) F_m and (b) weighted triadic pattern detection for the same commodity. while the fingerprint is typical in terms of absolute measures, the profile given after discounting reciprocity via null models is not trivial. RBCM signals two motifs, namely $m = 9$ and $m = 10$, i.e. the looped triad and its modified version with a bi-directional link. Instead, in weighted terms, RBCM+CRWCM signals only motif for $m = 1$, i.e. the open triad with one supplier and two users. For this commodity, a triadic type such as $m = 1$, that is not a motif in binary terms it is significatively so in weighted terms.

The triadic measures N_m and F_m in Fig.4.13(a) and (c) describe the predominance of $m = 1$, i.e. formation with one supplier connected to two users, and the almost totally abscence of the other type of triadic formations.

While the empirical values of occurrences and fluxes is as trivially expected formed by a high number of (and weight on) open triads, such as $m = 1$, the z-score profiles obtained when comparing to the statistical null model are not trivial. Specifically in Fig.4.13(b) various motifs and anti-motifs are shown, namely motif $m = 9$ and $m = 11$ and anti-motifs $m = 8$ and $m = 12$. Motif $m = 9$ has been described in terms of financial networks, and has been proven to be a significant actor in determining crisis, in fact it represents a cyclical loop, a formation that is vulnerable to attacks on each node. In weighted terms, instead, as shown in Fig.4.13(d), only $m = 9$ is signalled as motif having an unexpected amount of money on it. This result, further increase the vulnerability of micro-industries in this commodity layer. While it is not clear how an industry can be totally taken out of the picture, an exogeneous shock to a firm in an industry taking part in formation $m = 9$ can clearly result in a systemic propagation towards other industries because of cyclical relationships.

4.4.6 Tax Consultancy

Now let us discuss the commodity taking into account services in Tax Consultancy. This commodity layer consists of 68 suppliers and 822 users.



Top 5 industries for out-degree, in-degree, supply and use activity.

Top 5 Out-Degree		Top 5 In-Degree	
Description	Activity	Description	Activity
Accounting offices	7.8%	Accounting offices	0.4%
Law firms	7.7%	General government admin.	0.4%
Organizational consultancy firms	7.6%	Organizational consultancy firms	0.4%
Accountants-administration consultants	7.1%	Holdings (not financial)	0.4%
Tax consultants	7.1%	Accountants-administration consultants	0.4%
Top 5 Supply		**Top 5 Use**	
Description	Activity	Description	Activity
Accounting offices	33%	General government admin.	12.5%
Law firms	12%	Organizational consultancy firms	7.6%
Tax consultants	11%	Holdings (not financial)	5.2%
Other admin. offices	10.1%	Non-life insurance (excluding health insurance)	2.7%
Accountants-admin. consultants	9.9%	Advertising agencies	2%

Figure 4.14 (Top-Left) Counter-cumulative distribution of out-degrees and in-degrees for the different industries in the layer 'Legal & Accountancy: Tax Consultants', with (Top-Right) the top 5 ranked industries in terms of binary activity ($k_{i;out}/\sum_i k_{i;out}$ or $k_{i;in}/\sum_i k_{i;in}$). (Bottom-Left) Counter-cumulative distribution of supply and use volumes for the different industries in the layer 'Legal & Accountancy: Tax Consultants', with (Bottom-Right) the top 5 ranked industries in terms of supply/use activity ($s_{i;out}/\sum_i s_{i;out}$ or $s_{i;in}/\sum_i s_{i;in}$).

In Fig.4.14 the counter-cumulative distribution for out-degrees and in-degrees (on the top-left panel) and supply and use (bottom-left panel) are depicted. For 'out' statistics the distribution is plotted in blue, while the analogue for 'in' statistics are plotted in orange. In both cases, the statistics of interest is normalized by the sum across industries in the commodity layer.

In the top-left panel the in-degree distribution presents a very low activity for all industries, consider that the top 5 ranked industries show an activity of around 0.4%. The out-degree is instead very heterogeneous for median activity but homogeneous around the top 5, with an activity ranging from 7.1% to 7.8%. In both cases the leading industry is 'Accounting Offices' followed by 'Law Firms', for the number of users, and 'General Government Administration', for number of suppliers. The latter is the leading industry in use volumes with an activity of 12.5%, slightly above the second ranked "Organizational consultancy firms" with 7.6%. Contrarily, in terms of supply volumes there is higher heterogeneity in the top 5, with the leading industry 'Accounting Offices' taking 33% of activity with respect to the second ranked 'Law Firms' with 12%.

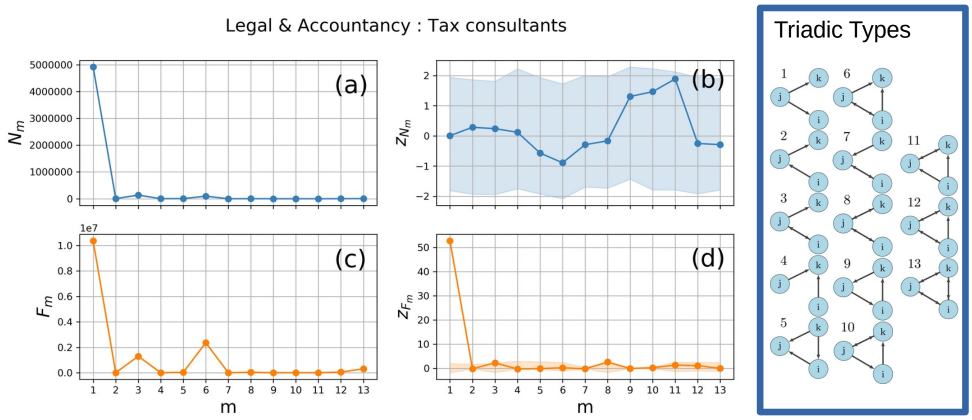


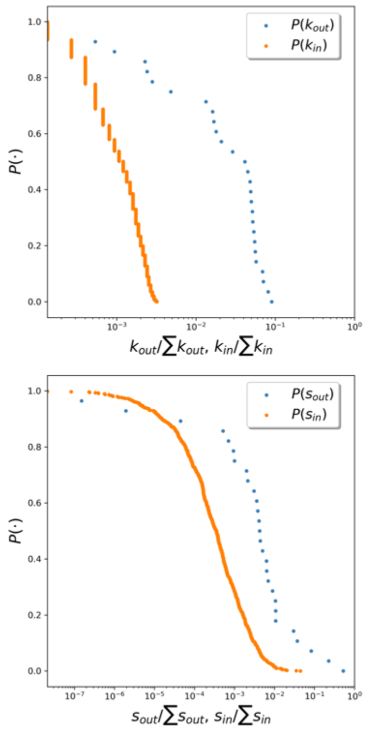
Figure 4.15 (a) N_m and (b) binary triadic pattern detection for the commodity 'Legal & Accountancy: Tax Consultants'. (c) F_m and (b) weighted triadic pattern detection for the same commodity. RBCM well describes the triads, no motifs are present when reciprocity of connections is taken into account. Instead, for what regards the total intensities of input/output relationships, $m = 1$ is a significative motif, i.e. an unexpected large amount of money lies on this type of triads according to the reciprocal model.

Fig.4.15 clearly shows a predominance of $m = 1$ both regarding occurrences (and hence N_m) and regarding triadic fluxes (and hence F_m). The measure for triadic fluxes also shows a relatively high concentration of money on $m = 3$ and $m = 6$.

When comparing empirical measures with statistical null models, no motifs are discovered for occurrences when taking into account both direction and reciprocity of links, as shown in Fig.4.15(b). Instead, an analysis on triadic fluxes, shows that $m = 1$ is clearly a weighted motif, i.e. an unexpectedly high amount of money is concentrated on triadic structures with one supplier and two users. The commodity layer, is hence, vulnerable to attacks on suppliers of Tax Consultancy services.

4.4.7 Electricity

Finally, we discuss the "Electricity" commodity layer, where industries are suppliers or users of such commodity. This layer consists of 28 suppliers and 797 users.



Top 5 industries for out-degree, in-degree, supply and use activity.

Top 5 Out-Degree		Top 5 In-Degree	
Description	Activity	Description	Activity
Prod. by thermal, nuclear and combined heat and power plants	9%	Supermarkets and similar stores	0.3%
Production of electricity through wind energy	8.2%	Restaurants	0.3%
Cultivation of apples and pears	7.1%	Shops in flowers and plants, seeds and garden supplies	0.3%
Production of electricity by solar cells, heat pumps and hydropower	6.9%	Fast food restaurants, cafeterias, ice cream parlors, food stalls, etc.	0.3%
Cultivation of other perennial crops	5.7%	Keeping dairy cattle	0.3%
Top 5 Supply		**Top 5 Use**	
Description	Activity	Description	Activity
Prod. by thermal, nuclear and combined heat and power plants	52.7%	Man. of tobacco products	4.5%
Prod. through wind energy	22.6%	General government admin.	3.5%
Prod. through solar cells, heat pumps and hydropower	8.3%	Supermarkets and similar stores	2.1%
Trade in electricity and gas through pipes	3.7%	Restaurants	1.9%
Cultivation of apples and pears	3%	University higher education	1.6%

Figure 4.16 (Top-Left) Counter-cumulative distribution of out-degrees and in-degrees for the different industries in the layer 'Electricity', with (Top-Right) the top 5 ranked industries in terms of binary activity ($k_{i;out}/\sum_i k_{i;out}$ or $k_{i;in}/\sum_i k_{i;in}$). (Bottom-Left) Counter-cumulative distribution of supply and use volumes for the different industries in the layer 'Electricity', with (Bottom-Right) the top 5 ranked industries in terms of supply/use activity ($s_{i;out}/\sum_i s_{i;out}$ or $s_{i;in}/\sum_i s_{i;in}$). A high concentration of supply volumes is around firms capable of producig electricity using thermal, nuclear, combined heat and power plants, namely the 52.7%. Instead use is scattered across a multiplicity of industries, leaded by 'Supermarkets' in number of suppliers and 'Manifacturing of Tobacco Products' in use volumes.

In Fig.4.16 the counter-cumulative distribution for out-degrees and in-degrees (on the top-left panel) and supply and use (bottom-left panel) are depicted. For 'out' statistics the distribution is plotted in blue, while the analogue for 'in' statistics are plotted in orange. In both cases, the statistics of interest is normalized by the sum across industries in the commodity layer.

'Electricity' is a utility, a commodity of first necessity and as such accommodates a large number of users connected to few suppliers, having the capability of producing this resource. The top 5 ranked industries for out-degree activity (number of users) are 'Production of electricity by thermal, nuclear and combinaed heat and power plants' with 9%, followed by 'Production of electricity through wind energy' with 8.2%. Understandingly enough also firms in 'Cultivation of apples and pears' are able to produce energy, reasonably from biomass, with an activity of 7.1%. While the out-degree activity in the top 5 is rather homogeneous, ranging from 5.7% to 9%, the same cannot be said about the supply volume activity ranging from 3% to 52.7%, with a clear supplying monopoly of firms producing electricity from thermal, nuclear, combined heat and power plants.

From the side of use, instead, the figure is relatively homogeneous. In-degree activity (number of suppliers) is aroud 0.3% for all of the industries in the top 5, namely 'Supermarkets', 'Restaurants', 'Flower Shops', 'Fast Food Restaurants and Cafeterias' and 'Keeping Dairy Cattle'. In terms of volumes, the top 5 have an activity ranging from 1.6% to 4.5% with as leading industry the 'Manifacturing of Tobacco Products'.

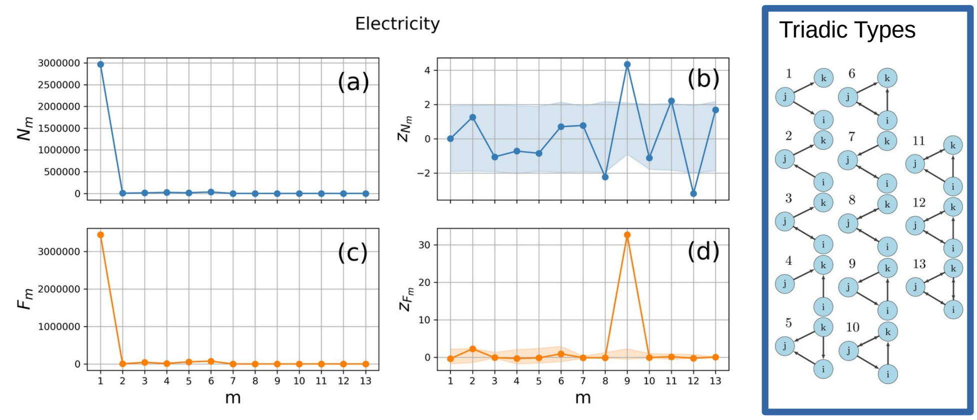


Figure 4.17 (a) N_m and (b) binary triadic pattern detection for the commodity 'Other Road Electricity'. (c) F_m and (b) weighted triadic pattern detection for the same commodity.

In Fig.4.17 we see the profile of values of N_m (on the left) and F_m (on the right) for 'Electricity'. Structure $m = 1$ is the triadic formation dominant both in terms of occurrences and in terms of triadic fluxes.

When passing to a description in terms of z-scores, we discover a collection of motifs and anti-motifs. Specifically, in binary terms in Fig.4.17 (left), $m = 8$ and $m = 12$ are anti-motifs while $m = 9$ and $m = 11$ are motifs. Even if $m = 9$ is not overly present in the data (almost absent), its value is unexpectedly high, i.e. also in electricity it exists a problem in case of bankrupcy of a single industry. More so, the z-score profile for triadic fluxes in Fig.4.17(Right) shows a motif $m = 9$, so the commodity layer in 'Electricity' is twofold vulnerable: industries in $m = 9$ structures can be broken when a single industry goes bankrupt (motif in binary terms), i.e. has a huge economic crash, or even when it has a significant reduction in input/output (motif in weighted terms).

5 Discussion

From the point of view of NSI's in general there are two main reasons to the systematic study of reconstruction techniques for networks. The first reason is that in many countries the necessary data at microlevel is not available to be able to track supplier-user relationships. As noted in the introduction, for countries where VAT data is registered per transaction between firms the data can be a great resource, although even there, a substantial data-cleaning effort is still required. However, for the Netherlands there is no such system unless a transaction involves cross-border traffic.

For a modest sample of companies supply-use relationships are collected using external datasuppliers, and totals in terms of turnover per firm are also known. These are certainly important constraints but they do not provide the necessary detail by themselves. Purely from the point of view of traditional National Accounting, the turnovers might be considered to be sufficient. However, there is an increasing awareness that National Accounts are inadequate for economic planning, for economic risk assessment, and for supply chain accountability in the context of environmental footprints, money laundering, or exploitation of labour. In all of these policy areas, there is a demand for reliable and comprehensive facts to support the decision- and policy-making process. All of these purposes require a much more fine-grained view of supplier-user relationships.

The ideal situation does remain to be that a ground truth of microlevel data is available, preferably at the level of the volume and/or value of exchange of goods or services between firms. Currently, the deterministic modelling that is implemented at Statistics Netherlands is a form of model-based (mass) imputation. A set of stylized 'rules' has been compiled, using expert economic insights and also empirically established correlations using as examples countries where transactional (VAT) data are available. As with imputation on a large scale in other socio-economic subject areas, there are reservations in using this for general-purpose statistical output because for any particular use the quality of the output might well suffer from biases that are poorly understood.

At the very least, as has been pointed out in previous CBS discussion papers on network reconstruction (Rachkov et al. 2021, Kayzel and Pijpers 2023), the deterministic reconstruction must continue to be confronted and contrasted with the results from a very different approach. Also, since the output of any reconstruction method is ultimately a model of the real (unobserved) network, from the point of view of goodness of fit it is of course crucial that the real network is given positive likelihood by the reconstruction method itself. This has vanishing probability to happen via a deterministic method, while probabilistic methods that admit all possible networks as outcomes automatically ensure a positive likelihood for the real one. Clearly, the challenge for probabilistic methods is that of producing informative probability distributions that are sufficiently concentrated around the real network. As is demonstrated in this paper, the unbiased maximum-entropy reconstruction method is perfect for this purpose, once it uses informative constraints as input.

Compared with the recent CBS papers (Rachkov et al. 2021, Kayzel and Pijpers 2023) one very important extension is introduced in the modelling. Explicit account is taken not only in the direction of the supply-use relationship, but also whether there is a reciprocity in such relationships. Especially for a reconstruction per commodity layer but between industries, rather than individual firms, as is done here, the expectation is that there is substantial reciprocity, which in fact should increase for increasing aggregation towards entire sector level. This is sufficiently robust for networks that advantage can be taken of the fact in adding further constraints to the reconstruction, and hence construct more precise and accurate network realisations that are genuinely closer to the true network.

In addition to the improvement in accuracy of the reconstruction, leading to smaller margins of uncertainty on any statistical measure of network properties, there is a benefit in using the network null model to identify empirical deviations from it, gaining an understanding of unexpected structural properties.

From the previous section, and also the tabular summary of results for each commodity layer shown in appendix C, it is clear that for many commodity layers there is no significant deviation at all from the null model. In by far the most commodity layers there are at most a few motifs; i.e. significant deviations from the null model for a few types of triads. This demonstrates that as a framework for structure of industry-to-industry relationships at the SBI5 classification level, the null model is indeed useful and informative. As one would expect, there are nevertheless also a substantial number of commodity layers that have a distinctive pattern of motifs. This is an indication that the structuring of these layers is distinctive and require further investigation. There are multiple potential causes that could underlie such structures. For instance, it may be that the structure of these layers is different because there is a very commodity specific use that imposes that structure. On the other hand, in analogy with the financial markets where this model has previously been applied, it can indicate that there are structural inefficiencies or vulnerabilities in the industries where that class of commodities play a central role. Perhaps, they could indicate frictions between legislative and market forces, or interactions with international firm-to-firm trade, for example. It is precisely the fingerprint of motifs for those commodity layers that help in identifying which industries are most of interest and why. Hence the network reconstruction technique is a very important addition to the analysis toolbox for structural economists.

Over the past few years there has been an increasing interest from the point of view of environmental policy, sustainability, and also exploitation of labour to understand the interdependencies in manufacturing and the combined footprints of goods and services supplied to end-users. In the absence of a full ground truth it is important to have reliable reconstruction methods that can, at least in a statistical sense, map out the most likely footprint in any of these aspects for any given commodity. By monitoring the network structure it also becomes feasible to monitor the effects that for instance the "Green Deal" policies are having on these structures and hence help in judging how effective those policy measures are and which (unintended) side effects become visible through trade relationships. In this sense a suite of much more targeted statistical indicators comes within reach, that illuminate how the economy *functions* rather than merely describing (collectively) the properties of the actors: static snapshots of what is intrinsically a dynamic process.

As a next step for Statistics Netherlands, the reconstruction techniques are to be made available as tools to the economic division to assist in targeting detailed analysis of the structure of the economy. Eventually, the combination of expert knowledge and the industry-to-industry (and potentially firm-to-firm) network motifs will provide the link between the motifs and the functioning in terms of resilience, redundancy, competition, and circularity, to name but a few, for all Dutch industry sectors. This might lead to enabling using the motif-fingerprint directly as an indicator for which policy risk or benefits are being incurred, and whether and what particular policy might be most beneficial to each in order to reach societal or policy goals.

6 Conclusion

The main contribution of this paper is in a comprehensive mapping of triadic motifs per industrial sector in the Netherlands, using appropriate null models for the interfirm network as outlined in the paper Di Vece et al. 2024. Triadic motifs are the smallest interconnected building blocks of complex networks, such as production networks. They can be detected as over-occurrences with respect to null models that only consider pairwise interactions.

The results of this paper show that, while the aggregate industry-to-industry network (where industries are linked by trade in any possible commodity) exhibits so many deviating triadic motifs that it would be impossible to use the reconstruction method to any degree of reliability, the individual commodity-specific layers of the industry-to-industry network show either no deviating motifs or at most a few of them. This basically indicates that different commodities can be characterized in terms of very simple fingerprints of deviation (if any) from the null model. The detected existence of significant motifs in some, but certainly not all, industrial sectors can be economically important. If they act as amplifiers of shocks and hence increase vulnerabilities but also if they in fact act to dampen out shocks and increase resilience. It could even be that some motifs, in combination with production delays, could act a endogenous generators of economic 'noise' or shocks (see e.g. Moran et al. 2024). The comprehensive sectorial analysis presented here identifies some sectors of particular interest, which should be the focus of such research.

It is worthwhile using the reconstructed network as a starting point, for instance for agent based modelling to test out the behaviour of the economic system under supply or demand shocks or random endogenous delays in production chains. Such research would then also identify which indicators or aggregates of network properties would be most suitable for policy makers as they would be economically the most influential and distinctive for the performance of the system.

References

Acemoglu, D. et al. (2012). "The Network Origins of Aggregate Fluctuations". en. In: *Econometrica* 80.5, pp. 1977–2016. ISSN: 1468-0262. DOI: 10.3982/ECTA9623.

Alfaro-Ureña, A. et al. (2018). "The costa rican production network: Stylized facts". In: *Research Paper Series*.

Asikainen, A. et al. (May 2020). "Cumulative effects of triadic closure and homophily in social networks". In: *Science Advances* 6.19, eaax7310. DOI: 10.1126/sciadv.aax7310.

Atalay, E. et al. (Mar. 2011). "Network structure of production". In: *Proceedings of the National Academy of Sciences* 108.13, pp. 5199–5202. DOI: 10.1073/pnas.1015564108.

Bacilieri, A. et al. (2023). *Firm-level production networks: what do we (really) know?* Unpublished.

Bardoscia, M. et al. (June 2021). "The physics of financial networks". en. In: *Nature Reviews Physics* 3.7, pp. 490–507. ISSN: 2522-5820. DOI: 10.1038/s42254-021-00322-5. URL: https://www.nature.com/articles/s42254-021-00322-5 (visited on 03/06/2023).

Buiten, G. et al. (Dec. 2021). *Reconstruction method for the Dutch interfirm network including a breakdown by commodity for 2018 and 2019 (v1.0)*. DOI: 10.13140/RG.2.2.16685.77286.

Cardoza, M. et al. (Sept. 2020). *Worker Mobility and Domestic Production Networks*. IMF Working Papers 2020/205. International Monetary Fund. URL: https://ideas.repec.org/p/imf/imfwpa/2020-205.html.

Carvalho, V. M. and A. Tahbaz-Salehi (2019). "Production Networks: A Primer". In: *Annual Review of Economics* 11.1, pp. 635–663. DOI: 10.1146/annurev-economics-080218-030212.

Carvalho, V. M. et al. (Dec. 2020). "Supply Chain Disruptions: Evidence from the Great East Japan Earthquake". In: *The Quarterly Journal of Economics* 136.2, pp. 1255–1321. ISSN: 0033-5533. DOI: 10.1093/qje/qjaa044.

Chacha, P. W., B. Kirui, and V. Wiedemann (2022). *Mapping Kenya's Production Network Transaction by Transaction*. DOI: https://dx.doi.org/10.2139/ssrn.4315810.

Chong, S. et al. (2019). "The role of small- and medium-sized enterprises in the Dutch economy: an analysis using an extended supply and use table". In: *Economic Structures* 8, p. 8. DOI: https://doi.org/10.1186/s40008-019-0139-1.

Cimini, G., R. Mastrandrea, and T. Squartini (2021). *Reconstructing Networks*. Elements in the Structure and Dynamics of Complex Networks. Cambridge University Press. DOI: 10.1017/9781108771030.

Cimini, G. et al. (Jan. 2019). "The statistical physics of real-world networks". en. In: *Nature Reviews Physics* 1.1, pp. 58–71. ISSN: 2522-5820. DOI: 10.1038/s42254-018-0002-6. URL: https://www.nature.com/articles/s42254-018-0002-6 (visited on 03/06/2023).

Colomer-de-Simón, M. et al. (Aug. 2013). “Deciphering the global organization of clustering in real complex networks”. en. In: *Sci Rep* 3.1, p. 2517. ISSN: 2045-2322. DOI: 10.1038/srep02517.

Demir, B. et al. (Jan. 2022). “Financial Constraints and Propagation of Shocks in Production Networks”. In: *The Review of Economics and Statistics*, pp. 1–46. ISSN: 0034-6535. DOI: 10.1162/rest_a_01162. eprint: https://direct.mit.edu/rest/article-pdf/doi/10.1162/rest_a_01162/1983885/rest_a_01162.pdf. URL: https://doi.org/10.1162/rest_a_01162.

Dhyne, E., G. Magerman, and S. Rubínova (2015). *The Belgian production network 2002-2012*. eng. Working Paper 288. NBB Working Paper.

Di Vece, M., D. Garlaschelli, and T. Squartini (2023a). “Deterministic, quenched, and annealed parameter estimation for heterogeneous network models”. In: *Phys. Rev. E* 108 (5), p. 054301. DOI: 10.1103/PhysRevE.108.054301. URL: https://link.aps.org/doi/10.1103/PhysRevE.108.054301.

Di Vece, M., F. P. Pijpers, and D. Garlaschelli (2023b). *Commodity-specific triads in the Dutch inter-industry production network*. arXiv: 2305.12179 [physics.soc-ph].

– (2024). “Commodity-specific triads in the Dutch inter-industry production network”. In: *Nature Scientific Reports* 14, p. 3625. DOI: 10.1038/s41598-024-53655-3.

Diem, C. et al. (May 2022). “Quantifying firm-level economic systemic risk from nation-wide supply networks”. eng. In: *Sci Rep* 12.1, p. 7719. ISSN: 2045-2322. DOI: 10.1038/s41598-022-11522-z.

Garlaschelli, D. and M. I. Loffredo (Oct. 2004). “Fitness-Dependent Topological Properties of the World Trade Web”. en. In: *Physical Review Letters* 93.18, p. 188701. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.93.188701. URL: https://link.aps.org/doi/10.1103/PhysRevLett.93.188701 (visited on 03/06/2023).

– (July 2008). “Maximum likelihood: Extracting unbiased information from complex networks”. In: *Phys. Rev. E* 78.1. Publisher: American Physical Society, p. 015101. DOI: 10.1103/PhysRevE.78.015101. URL: https://link.aps.org/doi/10.1103/PhysRevE.78.015101.

Hooijmaaijers, S. and G. Buiten (2019). “A methodology for estimating the Dutch interfirm trade network, including a breakdown by commodity”. In: *OECD Conference: New Approaches to Economic Challenges*, p. 18.

Jamakovic, A. et al. (Aug. 2009). “How small are building blocks of complex networks”. en. In: DOI: 10.48550/arXiv.0908.1143.

Jaynes, E. T. (May 1957a). “Information Theory and Statistical Mechanics”. en. In: *Physical Review* 106.4, pp. 620–630. ISSN: 0031-899X. DOI: 10.1103/PhysRev.106.620. URL: https://link.aps.org/doi/10.1103/PhysRev.106.620 (visited on 03/06/2023).

– (Oct. 1957b). “Information Theory and Statistical Mechanics. II”. en. In: *Physical Review* 108.2, pp. 171–190. ISSN: 0031-899X. DOI: 10.1103/PhysRev.108.171. URL: https://link.aps.org/doi/10.1103/PhysRev.108.171 (visited on 03/06/2023).

– (1982). “On the rationale of maximum-entropy methods”. In: *Proceedings of the IEEE* 70.9, pp. 939–952. ISSN: 0018-9219. DOI: 10.1109/PROC.1982.12425. URL: http://ieeexplore.ieee.org/document/1456693/ (visited on 03/06/2023).

Kayzel, J and F. P. Pijpers (June 2023). *Stochastically reconstructing business production networks using maximum entropy*. DOI: 10.13140/RG.2.2.12361.11362.

Kumar, A. et al. (2021). “Distress propagation on production networks: Coarse-graining and modularity of linkages”. In: *Physica A: Statistical Mechanics and its Applications* 568, p. 125714. ISSN: 0378-4371. DOI: https://doi.org/10.1016/j.physa.2020.125714. URL: https://www.sciencedirect.com/science/article/pii/S0378437120310128.

König, M.D. et al. (Sept. 2022). “Aggregate fluctuations in adaptive production networks”. In: *Proceedings of the National Academy of Sciences* 119.38, e2203730119. DOI: 10.1073/pnas.2203730119.

Maluck, J. and R.V. Donner (July 2015). “A Network of Networks Perspective on Global Trade”. en. In: *PLOS ONE* 10.7, e0133310. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0133310.
Mattsson, C. E. S. et al. (2021). “Functional Structure in Production Networks”. In: *Frontiers in Big Data* 4. ISSN: 2624-909X.
McNerney, J. et al. (2022). “How production networks amplify economic growth”. In: *Proceedings of the National Academy of Sciences* 119.1. _eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.2106031118, e2106031118. DOI: 10.1073/pnas.2106031118.
Milo, R. et al. (Oct. 2002). “Network Motifs: Simple Building Blocks of Complex Networks”. In: *Science* 298.5594, pp. 824–827. DOI: 10.1126/science.298.5594.824.
Mizuno, T., W. Souma, and T. Watanabe (July 2014). “The Structure and Evolution of Buyer-Supplier Networks”. en. In: *PLOS ONE* 9.7, e100712. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0100712.
Moran, J. et al. (2024). “Timeliness criticality in complex systems”. In: *Nature Physics*. DOI: 10.1038/s41567-024-02525-w.
Mungo, L. et al. (2023). “Reconstructing production networks using machine learning”. In: *Journal of Economic Dynamics and Control* 148, p. 104607. ISSN: 0165-1889. DOI: https://doi.org/10.1016/j.jedc.2023.104607. URL: https://www.sciencedirect.com/science/article/pii/S0165188923000131.
Ohnishi, T., H. Takayasu, and M. Takayasu (Dec. 2010). “Network motifs in an inter-firm network”. en. In: *J Econ Interact Coord* 5.2, pp. 171–180. ISSN: 1860-711X, 1860-7128. DOI: 10.1007/s11403-010-0066-6.
Peydro, J.-L. et al. (2020). *Production and financial networks in interplay: Crisis evidence from supplier-customer and credit registers*. URL: https://cepr.org/publications/dp15277.
Picciolo, F. et al. (June 2022). “Weighted network motifs as random walk patterns”. en. In: *New J. Phys.* 24.5, p. 053056. ISSN: 1367-2630. DOI: 10.1088/1367-2630/ac6f75.
Rachkov, A., F. P. Pijpers, and D Garlaschelli (Mar. 2021). *Potential Biases in Network Reconstruction Methods Not Maximizing Entropy*. DOI: 10.13140/RG.2.2.31861.29925.
Rungi, A., G. Morrison, and F. Pammolli (2017). *Global Ownership and Corporate Control Networks*. DOI: http://dx.doi.org/10.2139/ssrn.3031955. IMT Lucca EIC WP Series 07/2017: 07/2017.
Spray, J. and S. Wolf (Nov. 2018). “17 Industries without Smokestacks in Uganda and Rwanda”. In: *Industries without Smokestacks: Industrialization in Africa Reconsidered*. Oxford University Press. ISBN: 9780198821885. DOI: 10.1093/oso/9780198821885.003.0017. URL: https://doi.org/10.1093/oso/9780198821885.003.0017.
Squartini, T., G. Fagiolo, and D. Garlaschelli (Oct. 2011a). “Randomizing world trade. I. A binary network analysis”. en. In: *Physical Review E* 84.4, p. 046117. ISSN: 1539-3755, 1550-2376. DOI: 10.1103/PhysRevE.84.046117. URL: https://link.aps.org/doi/10.1103/PhysRevE.84.046117 (visited on 03/06/2023).
– (Oct. 2011b). “Randomizing world trade. II. A weighted network analysis”. en. In: *Physical Review E* 84.4, p. 046118. ISSN: 1539-3755, 1550-2376. DOI: 10.1103/PhysRevE.84.046118. URL: https://link.aps.org/doi/10.1103/PhysRevE.84.046118 (visited on 03/06/2023).
Squartini, T. and D. Garlaschelli (Aug. 2011). “Analytical maximum-likelihood method to detect patterns in real networks”. en. In: *New J. Phys.* 13.8, p. 083001. ISSN: 1367-2630. DOI: 10.1088/1367-2630/13/8/083001.
– (Jan. 2017). *Maximum-Entropy Networks: Pattern Detection, Network Reconstruction and Graph Combinatorics*. ISBN: 978-3-319-69436-8. DOI: 10.1007/978-3-319-69438-2.
Squartini, T., I. van Lelyveld, and D. Garlaschelli (Nov. 2013). “Early-warning signals of topological collapse in interbank networks”. en. In: *Sci Rep* 3.1, p. 3357. ISSN: 2045-2322. DOI: 10.1038/srep03357.

Vallarano, N., C.J. Tessone, and T. Squartini (2020). “Bitcoin Transaction Networks: An Overview of Recent Results”. In: *Frontiers in Physics* 8. ISSN: 2296-424X. DOI: 10.3389/fphy.2020.00286.
Vries, G. de et al. (2021). “Firm productivity and functional specialisation”. In: *The world economy* 44 (5), pp. 1232–1260. DOI: 10.1111/twec.13055.
Wang, Z. et al. (2022). “Motif Transition Intensity: A Novel Network-Based Early Warning Indicator for Financial Crises”. In: *Frontiers in Physics* 9. ISSN: 2296-424X.

Appendix A MLE for Binary Null Models

Let us tackle the problem of Maximum Likelihood Estimation for the RBCM. Its distribution for the adjacency matrix is given by Eq.24. We define the Log-Likelihood $\mathcal{L}_b = \ln(P(A))$, which after simple algebra can be rewritten as

$$\mathcal{L}_b = -\sum_i (\theta_i^{\rightarrow} k_i^{\rightarrow} + \theta_i^{\leftarrow} k_i^{\leftarrow} + \theta_i^{\leftrightarrow} k_i^{\leftrightarrow}) - \sum_{i,j<i} \ln(1 + x_i^{\rightarrow} x_j^{\leftarrow} + x_i^{\leftarrow} x_j^{\rightarrow} + x_i^{\leftrightarrow} x_j^{\leftrightarrow}), \tag{A.1}$$

where the $x_i^{\cdot}$ represent the exponentiated version of the Lagrange parameters, as already discussed in the main paper.

The first-order equations needed to estimate the set of $3N$ parameters $\{\theta_i^{\rightarrow}, \theta_i^{\leftarrow}, \theta_i^{\leftrightarrow}\}$ is given by the partial derivatives of $\mathcal{L}_b$ onto the mentioned parameters:

$$\begin{cases} \partial_{\theta_i^{\rightarrow}} \mathcal{L}_b &= -k_i^{\rightarrow,*} + \langle k_i^{\rightarrow} \rangle \\ \partial_{\theta_i^{\leftarrow}} \mathcal{L}_b &= -k_i^{\leftarrow,*} + \langle k_i^{\leftarrow} \rangle \\ \partial_{\theta_i^{\leftrightarrow}} \mathcal{L}_b &= -k_i^{\leftrightarrow,*} + \langle k_i^{\leftrightarrow} \rangle, \end{cases} \tag{A.2}$$

equating the empirical set of non-reciprocated and reciprocated degrees to their expected values.

The model is optimized using the Newton-Raphson method implemented in the python package NuMeTriS available on Pypi and GitHub.

Appendix B GLE for Weighted Null Models

Let us consider now the problem of Maximum Generalized-Loglikelihood Estimation (GLE) for the CRWCM. As already stated in the main paper it is possible to factorize the conditional weight distribution $Q(W|A)$ in a reciprocated $Q(W|A^{\rightarrow})$ and a non-reciprocated part $Q(W|A^{\leftrightarrow})$. The model assumes that we know the type of the connection established between each dyad, i.e. we know A completely.

For convenience, we treat here only the problem for the non-reciprocated component, the reciprocated one follows the exact same principles.

The function $Q(W|A^{\rightarrow})$ follows equation Eq.35 and hence the relative log-likelihood $\mathcal{L}_w = \ln(Q(W|A^{\rightarrow}))$ can be written as

$$\mathcal{L}_w = -\sum_i (\beta_i^{\rightarrow} s_i^{\rightarrow} + \beta_i^{\leftarrow} s_i^{\leftarrow}) + \sum_{i,j\neq i} a_{ij}^{\rightarrow} \left(\beta_i^{\rightarrow} + \beta_j^{\leftarrow}\right), \tag{B.1}$$

a term including the constraints weighted by their Lagrange parameters and minus the logarithm of the 'partition function' Z_A, describing how probabilities are partitioned in all available states. A parameter estimation done on this functional restricts the available configuration as those compatible with the specific empirical realization A^*. If, instead, we want our estimation to be consistent with the binary model, allowing for the model-dependent noise, we can substitute the connection probability to the empirical adjacency matrix component, i.e. $a_{ij}^{\rightarrow} \rightarrow (p_{ij}^{\rightarrow})_{model}$, defining a Generalized Log-Likelihood $\mathcal{G}_w$

$$\mathcal{G}_w = -\sum_i \left(\beta_i^{\rightarrow} s_i^{\rightarrow} + \beta_i^{\leftarrow} s_i^{\leftarrow}\right) + \sum_{i,j\neq i} p_{ij}^{\rightarrow} \left(\beta_i^{\rightarrow} + \beta_j^{\leftarrow}\right). \tag{B.2}$$

The use of GLE requires the choice of a binary model to which the CRWCM is estimated on. In order to take into account reciprocation of links, a natural choice is to mix CRWCM with RBCM.

The estimation of the set of $2N$ parameters $\{\beta_i^{\rightarrow}, \beta_i^{\leftarrow},\}$ follows by deriving $\mathcal{G}_w$ on them

$$\begin{cases} \partial_{\beta_i^{\rightarrow}} \mathcal{G}_w & = -s_i^{\rightarrow,*} + \langle s_i^{\rightarrow} \rangle \\ \partial_{\beta_i^{\leftarrow}} \mathcal{G}_w & = -s_i^{\leftarrow,*} + \langle s_i^{\leftarrow} \rangle \end{cases} \tag{B.3}$$

where $s_i^{\cdot,*}$ are the empirical non-reciprocated strength sequences and $\langle s_i^{\cdot} \rangle$ are the related ensemble average according to the RBCM+CRWCM, e.g.

$$\begin{cases} \langle s_i^{\rightarrow} \rangle & = \sum_{j\neq i} \dfrac{(p_{ij}^{\rightarrow})_{RBCM}}{\left(\beta_i^{\rightarrow} + \beta_j^{\leftarrow}\right)} \\ \langle s_i^{\leftarrow} \rangle & = \sum_{j\neq i} \dfrac{(p_{ij}^{\leftarrow})_{RBCM}}{\left(\beta_i^{\leftarrow} + \beta_j^{\rightarrow}\right)}. \end{cases} \tag{B.4}$$

The GLE estimation of the RBCM+CRWCM is available on the NuMeTriS package, using the Newton-Raphson method, on Pypi and GitHub.

Appendix C Summary of Pattern Detection Analysis

In the following table each commodity layer is described in terms of binary motifs (anti-motifs) and weighted motifs (anti-motifs). For each m triadic subgraph-type two symbols are depicted, the first for binary motifs and the second for weighted motifs. If a motif (higher occurrence or concentration of money than expected) is present, it is represented by a '+' symbol. If an anti-motif (lower occurrence or lower concentration of money than expected) is present, it is represented by a '-'. If no motifs or anti-motifs are present, than this event is represented by a dot '·'. For example, the commodity 'Cereals' is characterized by no motifs for all m triadic types, and is hence characterized by a double dot, respectively for binary and weighted motifs, for each type. Instead, the third commodity in the table, i.e. 'Potatoes', is characterized by a weighted motif for $m = 2$, represented by the symbol '·+' (no binary motif and a weighted motif), a weighted anti-motif in $m = 5$, represented by the symbol '· -', and likewise for $m = 12$.

Motif profile													
Commodity	1	2	3	4	5	6	7	8	9	10	11	12	13
Cereals	··	··	··	··	··	··	··	··	··	··	··	··	··
Oil-bearing Seeds	··	··	··	··	··	··	··	··	··	··	··	··	··
Potatoes	··	·+	··	··	·-	··	··	··	··	··	··	·-	··
Seeds	··	··	··	··	··	··	··	··	··	··	··	··	··
Vegetables	··	··	··	··	··	··	··	··	··	··	··	··	··
Flowers	··	··	··	··	··	··	··	··	··	··	··	··	··
Other Arable Crops	··	··	··	··	··	··	··	··	··	··	··	··	··
Fruit	··	··	··	··	··	··	··	··	··	··	··	··	··
Bulb/Plant/Cutting	··	··	··	··	··	··	··	··	··	··	··	··	··
Beef/Veal	··	··	··	··	··	··	··	··	··	··	··	··	··
Raw Milk	··	··	··	··	··	··	··	··	··	··	··	··	··
Pigs/Piglets	··	·+	··	··	··	··	··	-·	··	··	··	+·	··
Poultry	··	··	··	··	··	··	··	··	··	··	··	··	··
Eggs	··	··	··	··	··	··	··	··	··	··	··	··	··
Other An. Products	··	··	··	··	··	··	··	··	··	··	··	··	··
Agr. Services	··	··	··	··	··	··	··	··	··	··	··	··	··
Forestry Prods.	··	··	··	-·	··	··	··	··	··	··	··	··	··
Fresh Fish	··	··	··	··	··	··	··	··	··	··	··	··	··
Petrolium	··	··	··	··	··	··	··	··	··	··	··	··	··
Natural Gas	··	··	··	-·	··	··	··	··	··	··	··	··	··
Oth. Minerals	··	··	··	··	··	··	-·	-·	··	··	··	··	··
Mining Svcs.	··	··	··	··	··	··	··	··	··	··	··	··	··
Veal/Beef Meat	··	··	··	··	··	··	··	··	··	··	··	··	··
Pork	··	··	··	··	··	··	··	··	··	··	··	··	··
Poultry Meat	··	··	··	··	··	··	··	··	··	··	··	··	··
Oth. Meat	··	·+	··	··	··	··	··	··	··	++	··	··	··
Processed Fish	··	··	··	··	··	··	··	··	··	··	··	··	··
Potato Prods.	··	··	··	··	··	··	··	··	··	··	··	··	··
Fruit/Veg. Juice	··	··	+·	··	··	··	··	··	··	··	··	··	··
Fruit/Veg. prods.	··	··	··	··	··	··	··	··	··	··	··	··	··
Oils/Cattle	··	··	··	··	··	··	··	··	··	··	··	··	··
Edible Fats	··	··	··	··	··	··	··	··	··	··	··	··	··
Drinking Milk	··	··	··	··	··	··	··	··	··	··	··	··	··
Milk Powder	··	··	··	··	··	··	··	··	··	··	··	··	··
Butter/Butter Oil	··	··	··	··	··	··	··	··	··	··	··	··	··
Cheese	··	··	··	··	··	··	··	··	··	··	··	··	··
Oth. Dairy Prods.	··	++	··	··	··	··	··	··	··	··	··	··	··

Motif profile													
Commodity	1	2	3	4	5	6	7	8	9	10	11	12	13
Rice/Grain Prods.	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Starch Prods.	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Bakery Prods.	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	+·	· ·
Cocoa/Chocolate	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Coffee/Tea	· ·	·+	-·	· ·	+·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Oth. Food	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Animal Feed	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Wine/Liquors	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Beer/Malt	· ·	· ·	+·	· ·	· ·	-·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Non alcoholic beverages	·+	+·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Te·tile raw mat.	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Clothing	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Wood (plates)	· ·	· ·	· ·	·+	· ·	· ·	· ·	· ·	· ·	- -	· ·	· ·	· ·
Wood Prods.	· ·	· ·	· ·	· ·	· ·	· ·	· ·	- -	· ·	· ·	· ·	· ·	· ·
Paper/Cardboard/Pulp	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Oth. Paper/Cardboard Prods.	·+	+·	-·	· ·	-·	+·	· ·	-·	· ·	· ·	+·	-·	+·
Printing	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Oth. oil, coal prods.	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	·+	· ·	· ·
Gasoline	· ·	+·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Jet Fuel	· ·	· ·	· ·	-·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Diesel	· ·	++	· ·	· ·	-·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Heavy Fuel Oil	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Oth. Liquid Combustibles	· ·	++	· ·	· ·	-·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Chemical Processing: Raw Materials	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Organic Raw Materials	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Fertilizer	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Synthetic Resins	· ·	· ·	· ·	· ·	· ·	· ·	· ·	+·	· ·	+·	· ·	· ·	· ·
Pest Control	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Paint/Varnish/Ink	· ·	· ·	· ·	· ·	-·	· ·	· ·	· ·	· ·	· ·	·+	· ·	· ·
Soap/Polish/Perfume	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Oth. Chemical Prods.	·+	· ·	· ·	· ·	· ·	·+	· ·	· ·	· ·	· ·	++	· ·	· ·

	Motif profile												
Commodity	1	2	3	4	5	6	7	8	9	10	11	12	13
Pharmaceutical Raw Materials	·+	· ·	· ·	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Pharmaceutical Prods.	· ·	· ·	· ·	· ·	·+	· -	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Rubber Bands & Oth. Rubber Prods	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Plastic Sheeting	·+	· ·	· ·	·+	-+	+·	· ·	- ·	· ·	· ·	+·	· ·	· ·
Plastic Packaging	· ·	· ·	+·	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Plastic Construction Goods	· ·	·+	+·	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Oth. Plastic Prods.	·+	+·	· ·	· ·	- ·	+·	· ·	· ·	· ·	- ·	++	- -	· ·
Oth. Building Material	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	+·	· ·	· ·
Glass Prods.	· ·	+·	- -	· ·	- -	+·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Cement and Ceramic Prods.	·+	·+	· ·	·+	·+	·+	- ·	· ·	· ·	· ·	+·	· ·	· ·
Concrete Prods.	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Ferrous Sheeting & Tubing	· ·	· ·	· ·	· ·	· ·	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·
Non-Ferrous metals and prods.	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Oth. Metal Prods.	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	·+	· ·	· ·
Metal Components for Doors & Windows	· ·	· ·	· ·	· ·	· ·	+·	· ·	- ·	- ·	· ·	· ·	- ·	+·
Electrical Components	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Computer, Peripherals, Components	·+	++	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Transmitter/Telephone	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Radio/TV/Video	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Measurement Devices	·+	+·	· ·	· ·	· ·	+·	· ·	· ·	· ·	- ·	+·	- -	· ·
Photo Devices	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	+·	· ·	· ·	· ·
Electrical Devices	·+	++	· ·	+·	- ·	· ·	· ·	· ·	· ·	· ·	+·	· ·	· ·
Household Devices	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	+·	· ·
Oth. Machine Devices	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	+·	· ·	· ·	· ·	· ·

Motif profile													
Commodity	1	2	3	4	5	6	7	8	9	10	11	12	13
Oth. Car & Bus Semitrailers	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Auto Components	·+	· ·	· ·	· ·	- ·	- ·	· ·	+·	· ·	· ·	· ·	- ·	· ·
Ships/Floating Materials	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Planes & Components	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Motorcycles & Oth. Transport & Components	· ·	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·	· ·	++	· ·	· ·
Furniture & components	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	+·	· ·	· ·
Oth. Goods	·+	· ·	· ·	· ·	- ·	·+	· ·	· ·	· ·	· ·	+·	- ·	· ·
Repair, Installation, Maintenance	·+	· ·	·+	+·	· ·	·+	· ·	· ·	· ·	+·	· ·	· ·	· ·
Electricity	· ·	· ·	· ·	· ·	· ·	· ·	· ·	- ·	++	· ·	+·	- ·	· ·
Gas/Hot Water/City Heating	·+	· ·	·+	· ·	· ·	· ·	· ·	· ·	· ·	- ·	· ·	· ·	· ·
Net and Land Energy	· ·	· ·	· ·	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Water	· ·	++	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Environmental Service Government	· ·	· ·	· ·	· ·	- ·	+·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Environmental Service Commercial	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Cleaning & purification concessions	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Waste Metal/Slag	· ·	· ·	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Other & Dang. Waste	· ·	· ·	· ·	· ·	· ·	++	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Constr: New Housing	· ·	·+	· ·	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Constr: Building Maintenance	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Constr: Tunelling, Waterways, Roads	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Spec. Constr: New Housing	·+	·+	· ·	·+	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Spec. Constr: House Maintenance	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Spec. Constr: New Building	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	+·	- -	· ·	· ·
Spec. Constr: Building Maintenance	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Spec. Constr: Tunelling, Waterways, Roads	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·

Motif profile													
Commodity	1	2	3	4	5	6	7	8	9	10	11	12	13
Installation Insulation House	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Installation Insulation Other Buildings	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Installation Maint. Tunelling, Waterways, Roads	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Maintenance Cars & Motorcycles	·+	·+	· ·	· ·	· ·	· ·	+·	· ·	· ·	· ·	· ·	-·	· ·
Trading Services	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Train Transport Services	· ·	· ·	· ·	-·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Tram & Bus transport Services	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Oth. Road Transport & Pipelines	·+	· ·	· ·	· ·	-·	· ·	· ·	· ·	· ·	+·	· ·	· ·	· ·
Taxi Services/Oth. people transport	· ·	· ·	· ·	· ·	· ·	-·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Shipping Services	· ·	+·	-·	·+	--	· ·	-·	· ·	· ·	· ·	+·	-·	· ·
Airline Services	· ·	· ·	· ·	· ·	· ·	-·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Oth. Transport-related services	·+	· ·	· ·	·+	-·	+·	· ·	-·	-·	--	+·	· ·	+·
Postal Services	·+	++	+·	· ·	· ·	·-	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Courier Services	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Catering Hospitality Services	· ·	· ·	+·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Catering Food & Drink	· ·	· ·	· ·	· ·	· ·	· ·	· ·	-·	+·	· ·	· ·	· ·	· ·
(E-)Books/Calendars	· ·	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Ads	· ·	· ·	· ·	· ·	· ·	· ·	·-	· ·	· ·	· ·	+·	· ·	· ·
Magazine/Newspapers	· ·	· ·	+·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Printed Matter & Online	·+	·+	· ·	++	-+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Software/Games	·+	· ·	· ·	· ·	· ·	-·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Film/Cinema/Sound	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Music on Carrier or Paper	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Radio & TV: state & Commercial	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Telecomm. Services	· ·	· ·	· ·	·+	·+	· ·	·+	· ·	· ·	· ·	·+	· ·	· ·
Computer Services	· ·	· ·	· ·	·+	-·	· ·	·-	· ·	· ·	· ·	· ·	· ·	· ·

Motif profile													
Commodity	1	2	3	4	5	6	7	8	9	10	11	12	13
Web Services/News Agencies	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· -	· ·	· ·	· ·
Commission Financial Institution	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
FISIM	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Non-Life Insurance	· ·	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Financial Emergency Services	· ·	·+	· ·	· ·	+·	· ·	· ·	· ·	· ·	- -	· ·	- -	· ·
Estate Agents	· ·	· ·	+·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Rent Commercial Building	·+	· ·	- ·	- ·	- ·	+·	- ·	- ·	· ·	· ·	++	· ·	++
Rent Commercial Building (exempt)	· ·	· ·	- ·	- ·	· ·	+·	· ·	- ·	· ·	- ·	· ·	· ·	+·
Legal & Accountancy: Ta· Consultants	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Intra-Group Services	· ·	· ·	· ·	· ·	- ·	+·	· ·	· ·	· ·	·+	· ·	·+	·+
Economic & Management Consultancy	·+	· ·	· ·	· ·	· ·	· ·	· ·	·+	· ·	· ·	· ·	· ·	· ·
Engineer & Architecture & Building Inspection	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	+·	· ·	· ·
R & D	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Advertising, Market & Opinion Research	· ·	· ·	·+	·+	· ·	·+	·+	·+	- ·	- ·	·+	·+	·+
Specialist Business Services	·+	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·	· ·	++	· ·	· ·
Veterinary Services	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Car Rental & Leasing	·+	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·	· ·	++	· ·	+·
Licenses, Royalties, Agricultural	·+	· ·	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Rental of other real estate	·+	· ·	- ·	- ·	· ·	++	- ·	- ·	· ·	- ·	+·	· ·	++
Temporary Workforce agencies and intermediaries	·+	+·	- ·	- ·	- ·	+·	- ·	- ·	+·	· ·	·+	· ·	+·
Travel Agency	· ·	· ·	· ·	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Travel Organization	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Info/Booking	· ·	· ·	· ·	++	· ·	· ·	· ·	· ·	++	· ·	· ·	· ·	· ·

	Motif profile												
Commodity	1	2	3	4	5	6	7	8	9	10	11	12	13
Security & Detective Services	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Gardening, Landscaping, Cleaning Services	· ·	· ·	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·	+·	· ·	· ·
Oth. Commercial Services	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Health Care providers; government	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Governmental Market Services	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Education Services (agencies)	· ·	·+	· ·	·+	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Private/Commercial Education	· ·	· ·	· ·	- ·	· ·	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·
Hospital Care	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Oth. Health Care	·+	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Oth. Welfare Work	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Entertainment & Art	·+	- ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Cultural Institutions, Museums, Libraries	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Sports Recreation	· ·	· ·	· ·	· ·	- ·	· ·	· ·	· ·	· ·	· ·	++	· ·	· ·
Social & Charitable organisations	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Art conservation & repair	-+	+·	· ·	+·	- ·	+·	· ·	· ·	++	· ·	· ·	· ·	· ·
Personal Services	· ·	· ·	· ·	·+	· ·	· ·	· ·	· ·	· ·	+·	· ·	· ·	· ·
Intermediary Construction (taxed)	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·
Intermediary Construction (exempt)	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·	· ·

Colophon

Publisher
Statistics Netherlands
Henri Faasdreef 312, 2492 JP The Hague
www.cbs.nl

Prepress
Statistics Netherlands, Grafimedia

Design
Edenspiekermann

Information
Telephone +31 88 570 70 70, fax +31 70 337 59 94
Via contact form: www.cbs.nl/information