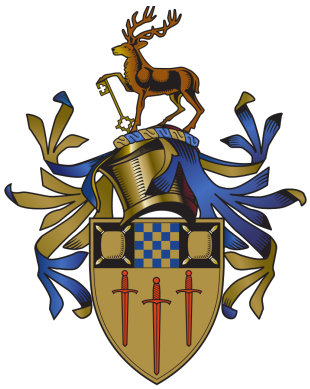
A Network Theory Approach to the Study of International Commodity Markets

Craig Robert Shenton

Thesis submitted to the University of Surrey
for the degree of Doctor of Philosophy

Centre for Environment and Sustainability
University of Surrey



Copyright © 2016 by Craig Robert Shenton. All rights reserved.
E-mail address: c.shenton@surrey.ac.uk

www.manaraa.com

ProQuest Number: 10624902

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 10624902

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

www.manaraa.com

Abstract

The international trade in commodities forms a complex network of economic interdependencies. This network now plays a central role in promoting global economic development and security. However, significant asymmetries have been noted in terms of access to this network, and in the unequal distribution of the benefits and risks accrued from the system as a whole. Understanding the statistical properties and dynamics of the trade network have therefore, become important tools for investigating a multitude of real-world policy concerns relevant to economics, public policy, and international development. This thesis focuses on investigating three of these issues—market growth, price inequality, and supply risks. The first of these projects focuses on modelling the growth of commodity markets, and the resulting effect on network topology. The second, looks at how asymmetries in network can lead to varying prices for the same good, and explores the implications for developing more equitable market structures. The final project contributes to our understanding of how export restrictions affect the network structure of trade and how these risks can undermine global food security. Throughout, a network science approach is employed, whereby trade is modelled as a graph-like structure, with the topology of trade being the primary focus of analysis. To support this approach, we introduce several theoretical models, and apply simulations on both real-world, and artificially produced trade network data. The outcome of this research improves on our ability to identify and target key participants within a market, and predict policies that favour more stable and equitable structures that better facilitate trade.

www.manaraa.com

Declaration

This thesis and the work to which it refers are the results of my own efforts. Any ideas, data, images or text resulting from the work of others (whether published or unpublished) are fully identified as such within the work and attributed to their originator in the text, bibliography or in footnotes. This thesis has not been submitted in whole or in part for any other academic degree or professional qualification. I agree that the University has the right to submit my work to the plagiarism detection service TurnitinUK for originality checks. Whether or not drafts have been so-assessed, the University reserves the right to require an electronic version of the final document (as submitted) for assessment as above.

Craig Robert Shenton
April 2017

www.manaraa.com

Acknowledgements

First and foremost, I must express my gratitude towards my supervisors, Prof Angela Druckman and Dr David Lloyd, who offered me invaluable advise and guidance throughout. Furthermore, I would like to extend thanks to Dr Spencer Thomas for his feedback on Chapter IV. In addition, I could not have completed my research without funding from the ERIE project, the University of Surrey Centre for Environmental Strategy, and the UK Engineering and Physical Sciences Research Council.

www.manaraa.com

Table of contents

www.manaraa.com

www.manaraa.com

www.manaraa.com

www.manaraa.com

List of figures

المنارة للاستشارات
www.manaraa.com

www.manaraa.com

List of tables

المنارة للاستشارات

www.manaraa.com

Chapter 1

Introduction

The past thirty years has been an unprecedented era of globalisation, with many developing economies rapidly integrating into a global trading system [139]. This system has formed together to create a complex network of economic interdependencies. In this context, markets facilitating the exchange of physical goods and commodities, have become a major source of income for many in the developing world, with a country's success in this arena, now a major predictor of overall economic performance [126]. However, significant asymmetries have been noted between countries that rely on this system, both in terms of access to the network, and in the unequal distribution of benefits accrued from the system as a whole [126]. As one of the most important interaction channels between countries, trade can also act as a conduit for spreading risks [122]. In highly interconnected markets, even small economic perturbations can reach across great distances and become truly global crises [44]. Disruptions in the energy and agriculture sectors in particular, can have non-trivial implications in terms of food availability and energy security—for both governments and businesses alike.

Given the central role played by the international trading network in promoting global economic development and security, the study of its statistical properties and dynamics, have become an issue of growing interest in the research community. Understanding the movement of goods across the network, and the interconnections that this activity creates, is also becoming integral in forming coherent international trade policy [81, 139]. Despite its recent emergence, the network approach to trade has produced many theoretical models and experimental tools necessary to quantify and measure market behaviour. However, far less progress has been made in connecting these statistical properties to real-world outcomes, which motivates the central aim of this thesis. In quantifying commodity markets in terms of a network topology, we can develop tools for investigating real-world policy concerns relevant to readers in economics,

www.manaraa.com

public policy, and international development.

While these concerns are too broad to be fully covered by any one thesis, we begin by addressing just three specific issues—*market growth*, *price inequality*, and *supply risks*—in a series of empirically based projects that are detailed below. Each project focuses on measuring and understanding one of these concepts in a network setting, but grounded in a theory of economic behaviour and interaction. Throughout the thesis, a *network science* approach is employed, whereby trade is modelled as a graph structure, with the topology of trade being the primary focus of analysis. To support this approach, we introduce several theoretical models, and apply simulations on both real-world, and artificially produced trade network data. This approach is important for two reasons. First, it will help the policy community understand how the structure of trade networks affects the behaviour and dynamics of commodity trading. Second, if we can predict policies that favour more stable and equitable structures that better facilitate trade, we can, in turn, aid in the restructuring of future trade networks. More broadly, this thesis also contributes to the exploration of the intersection between theories of trade and the literature on graph theory.

Before outlining the remainder of the thesis, we give some preliminaries on commodities trading, including a basic history and notable developments, along with a summery of how commodities are important to international trading as a whole. We also briefly outline the network science approach to modelling trade data that is used throughout the thesis, and discuss how and where it differs from standard economic methods.

1.1 Commodity Markets

A 'commodity' is defined as a good that is supplied without *qualitative* differentiation. That is to say, a specific commodity purchased from the U.S., should theoretically, be of the same quality, colour, flavour, as a commodity purchased from China. This characteristic of *substitutability*, greatly facilitates trade, as any importer can trade with any exporter and *vice versa*, knowing they are purchasing essentially the same good [128]. Commodities largely consist of 'raw' materials that have not been processed into other products. Consequently, many commodities are the inputs for a wide array of manufacturing processes. Commodities of this type—solid and liquid fuels, or industrial and precious metals for example—are referred to as 'hard' commodities. Soft commodities, on the other hand, are farm products such as, grains, food-stuffs, and plant fibres. A smaller grouping of these soft commodities, that are specific

www.manaraa.com

to the tropics, such as sugar, cocoa, and coffee beans, are also referred to as 'softs'. Table 1.1 details the major commodity types and list the primary exchanges on which these commodities are sold [27]. Note that the commodity exchanges listed only refer to where the contracts are traded—the physical goods are shipped directly from the originating country to the purchasing country.

Type	Commodity	Symbol	Exchange	Abbreviation
Energy	Crude Oil	CL	New York Mercantile Exchange	NYMEX
	Brent Crude	B	International Petroleum Exchange	IPE
	Ethanol	AC	Tokyo Commodity Exchange	TOCOM
	Natural Gas	NG	Central Japan Commodity Exchange	CJCE
Metals	Copper	HG	New York Mercantile Exchange	COMEX
	Lead	LME	London Metal Exchange	LME
	Aluminium	LMAHDS	Shanghai Futures Exchange	SFE
	Gold	GC	Philadelphia Board of Trade	PHLX
	Platinum	XPTUSD	Tokyo Commodity Exchange	TOCOM
Fibres	Wool	ASX	Chicago Mercantile Exchange	CME
	Cotton	CT	New York Cotton Exchange	NYCE
Grains	Maize	C	Chicago Board of Trade	CBT
	Rice	ZR	Dalian Commodity Exchange	DCE
	Soybeans	S	Kansas City Board of Trade	KCBT
	Wheat	EBL	Minneapolis Grain Exchange	MGE
Softs	Sugar	SB	Coffee, Sugar and Cocoa Exchange	CSCE
	Cocoa	CC	New York Board of Trade	NYBOT
	Coffee	KC	EURONEXT, UK	EURONEXT

Table 1.1 Globally traded commodities by type and by exchange typically listed.

The commodities listed in Table 1.1 account for approximately 23% of all world trade (measured in US$ worth of trade), with the remainder being trade in intermediate or manufactured products [91]. Commodities typically have little or no value added, are sold in bulk, and are traded by many countries, both in imports and exports, leading to highly competitive markets [91]. These features, along with relatively stable demand for what are every day essentials, lend themselves to a very low unit price. An important caveat being that while demand is relatively stable, both the short- and long-term supply of many commodities is often not. This leaves many markets with a high degree of price volatility (albeit starting from a low price point). The issue here, is that the time required to make changes in production is

www.manaraa.com

generally quite long, especially compared to manufacturing, in which output (i.e., supply) can be changed 'just-in-time' in some cases. For example, increasing the extraction of raw minerals can take years, or even decades if new mines have to be planned and built [91]. Even then, price volatility makes investment decisions about new production capabilities difficult, and often introduces price risks further down the supply chain. In the soft commodities, even small variation in annual yields can lead to price volatility, due to the extremely low (or zero) price 'elasticity' (i.e., the ability or willingness of buyers to accept higher prices) of many staple crops. In recent years, these risks have been further exacerbated by the renewed use of restrictive export policies in several important grain markets. The implications of these actions is discussed and analysed further in Chapter VI of the thesis.

Historically, the origins of modern commodity exchanges can be traced back to 17th century Holland, where agricultural products began to be bought and sold on formal contract. Commodity exchanges only became a major factor in international trade after the industrial revolution. The United Kingdom, in particular, became a major importer of industrial metals. The rapid increase in traded goods was organised formally as the London Metal Exchange (LME) in 1877. Energy commodity markets are a far more recent development, with the introduction of gas oil futures contracts on the International Petroleum Exchange (IPE) coming as recently as 1981. Soon after sweet crude oil contracts were listed on the New York Mercantile Exchange in 1983, and Brent crude futures in 1988 [91]. China's recent expansion into international trade has also greatly influenced the development of commodity trading. In 1994, the Chinese Securities Regulatory Committee consolidated all local exchanges into just three internationally focused exchanges; The Shanghai Futures Exchange (SFE), The Zhengzhou Commodity Exchange (ZCE), and The Dalian Commodity Exchange (DCE). After China's trade liberalisation efforts, commodity exchanges have proliferated across Asia, as more countries followed suit and begun to deregulate their economies. Despite this expansion, there remains a high degree of concentration, in terms of the value of trade exchanged on each market, with the majority focused in just the US, Japan, China, and the UK [27].

International trade policy on commodities is driven primarily by the World Trade Organisation (WTO), which sets rules (of which only some are legally binding) on the kinds of trade restrictions countries can put in place. The WTO was formally established in 1995, but its predecessor organisation, General Agreement on Tariffs and Trade (GATT) has been operating since 1948. The primary aim of the WTO is to dismantle barriers to trade, and has been a major influence in the globalisation of trade [131]. However, the WTO faces difficulties in balancing

www.manaraa.com

international trade policy between the developing and developed worlds. On the one hand, low commodity prices are typically favoured by developed industrial nations, as they are generally net importers (with notable exceptions including Australia and Canada). However, many net exporters are among the least developed nations in the world, and rely on selling commodities for the majority of their income. Naturally, they favour higher prices, that will increase export earnings and stimulate economic growth. Direct market interventions such as production limits, export quotas, and buffer stocks can affect prices one way or the other. However most policies of this nature have now been removed, largely due to the influence of the WTO's trade liberalisation efforts [91]. While nominally, countries cannot discriminate between trading partners under the WTO rules, many exceptions persist. For example, the Europeans Union's common market arrangement discriminates against goods imported from outside the EU. On the whole, however, far more discriminatory policies have been adopted by those countries that are still in the process of industrial development [131]. This differentiation is codified to some extent in the WTO rules under the principle of 'special and differential treatment', which has given developing countries the ability to impose tariffs on up to two-thirds of imports, and excluded them from the rules regarding non-tariff barriers (i.e., quantitative trade restrictions). However, it is thought that allowing developing nations to 'opt-out' of the WTO rules is only harming their economic prospects by creating an unfavourable investment climate that could discourage future trade and investments [131].

1.2 Modelling Trade as a Network

We now turn to discussing the network science approach to modelling commodity trade data. While economists have a long history of research on mathematical models of trade and markets, it has been rare for the network structure to be explicitly considered [121]. However, new approaches have recently emerged from outside the field. Influenced by developments in physics and computer science, researchers have begun applying graph-theoretic approach to modelling trade data [122, 43]. This approach, often referred to as *network science*[1], has attracted considerable attention due to the supposed universal properties of networks, with applications found across a wide range of disciplines [13]. The proponents of the network science approach argue that these new modelling techniques can reveal properties of trade which cannot be fully explained by standard economic modelling, as well as account for previously unexplained empirical observations [36, 122]. However, many of the more abstract network concepts, do not, as yet, have direct economic interpretations [85]. For this reason,

[1] Network science is also known as 'social network theory'.

www.manaraa.com

difficulties remain in assessing the relevance of many network properties and in connecting them to real-world outcomes [121].

At the most basic level, a network model of international trade consists of a number of vertices, each representing a particular country, and edges drawn between these vertices, representing bilateral trade in the form of imports and exports of a given commodity. A country's degree can then be defined as the number of edges to or from that country, with imports as the in-degree and exports as the out-degree [122]. Many countries trading with each other, in often divergent and asymmetric patterns, form together to give the network its structure. Where network science differs most from traditional economic modelling, is that it primarily focuses on measuring the topology of a given network and developing generative models that can reproduce networks of a similar structure. Due to the size and complexity of the international trade data, many economic models rely on aggregating many trade relationships into single variables [52]. However, using the network approach, researchers have been able to extract hitherto unknown statistical properties from these data [122]. More importantly, through network modelling, we can also consider relationships that go beyond the scope of standard international-trade indicators, which only account for direct bilateral-transactions. For example, second-order topological properties can be measured, such as, the level of indirect competition between exporters, and the formation of distinct communities within the network [44]. It is also thought that the global structure of the international trading system can indirectly influence the trading behaviour of individual countries, leading to another area of analysis [123].

Beyond descriptive statistics, researchers have also developed a wide array of generative models that may be useful in exploring the evolution and growth of the international trading system. The generative models have many characteristics in common. They typically begin with a small seed network, often of just a few countries (vertices) connected together at random. This is used to get the algorithm started. They then proceed to build the network sequentially, adding one country at a time. Every time a pair of countries makes a trade, an edge is introduced between their respective vertices. As the network grows, more and more vertices and edges are introduced, growing the size and density of the network accordingly. Usually new edges are added via a stochastic dynamic [132], with the 'preferential attachment' model in particular, being considered of importance to international trade [16, 122]. In this case, a new country entering the market is likely to trade with an existing country with a probability proportional to that existing country's degree [29]. This process is often thought to mimic how highly

www.manaraa.com

connected countries at the core of the network acquire trade relations at a faster rate than those nations on the periphery [29].

1.3 Thesis Structure

The remainder of this thesis is structured as follows;

CHAPTER II

An Introduction to Networks: The next chapter reviews some basic graph theory notation and concepts that are used throughout the thesis. Graph notation for directed, weighted, and bipartite networks are provided, and a comprehensive list of network measures commonly in use in the literature are explained.

CHAPTER III

Review of the Trade Network Literature: The third chapter provides a comprehensive review of the models and experiments developed thus far in the trade network literature, and gives an overview of the results and conclusions drawn from these works.

CHAPTER IV

Growth in Network Models of Commodity-Trade: A ubiquitous feature of the trade network literature are reports of extremely high levels of heterogeneity in the degree distribution of the networks under study. These results are often taken to imply a so-called 'preferential attachment' model of network growth. Recently however, serious critiques have been levelled at the methods used in these studies, and consequently, we need to re-evaluate many of the conclusions drawn. In this chapter, these critiques are outlined and the correct methods for measuring the degree distributions of empirical networks are identified. An alternative methodology is employed, that directly measures the growth in network connectivity. This shows that growth in many trade networks is constrained, resulting in a far different topology than the preferential model would imply. A new single parameter model is developed that reproduces these features across a wide array of commodity markets. This model has wide reaching implications, both in terms of estimating future market growth and the robustness of critical markets.

www.manaraa.com

CHAPTER V

Price Variation in Network Models of Commodity-Trade: Whilst the international trade in commodities is a major source of income for many developing countries, restrictions on trade introduce asymmetries in the economic interactions between consumers. These asymmetries can lead to varying prices for the same goods simply due to the network structure. While having to pay higher prices is detrimental for low-income consumers, low prices can prove challenging for producers as well. In this Chapter, a network model is applied to commodity markets to measure the effect of a network structure on both local and global price asymmetries. A simple economic model is used, whereby standard price formation factors are specified so that variation in prices can be ascribed solely to the network structure. Using the model, the local network determinants of price variation are investigated, along with how global prices vary between networks with different topological properties. These findings combine theoretical analysis, modelling, and applications to real-world trade data. The results are interpreted in terms of both economic theory and international trade policy.

CHAPTER VI

The Risk of Export Restrictions to Food-Commodities: Research has shown that export restrictions play a key role in exacerbating the risks faced by food-import dependant countries. These risks are evidenced by the recent episode of extreme food-price volatility, commonly referred to as "the global food crisis". From a network perspective, a country's susceptibility to these risks can be directly related to its level of integration into the world trade system's structure. In this chapter, a network theory approach is used to analyse the effect of export restrictions on three major food-commodity markets. This method allows for the extraction of relevant structural information about the effect of export restrictions on the network as a whole, whilst also identifying the relative risks to specific countries. A composite risk index is developed that assesses the risk posed by export restrictions to food-importing countries based on their position in the network. These risks are interpreted in terms of developing a policy of strategic reserves. A case study example is also provided, showing how these network metrics can be applied to improve food-security policies.

CHAPTER VII

Conclusion and Outlook: The concluding chapter provides an overview of all results from the three projects, and offers some discussion on the policy implications with regard to negotiating global trade agreements. This chapter also details the limitations of the methodologies employed, and outlines areas for further development.

المنارة للاستشارات
www.manaraa.com

Chapter 2

An Introduction to Networks

2.1 Notation

In this chapter, we review some basic graph theory. The definitions in this chapter are taken from Barabási [13], unless otherwise stated. The following notation is used throughout the thesis:

- $\mathbf{A}$ = Adjacency matrix of a graph
- a_{ij} = Indicator of an edge between vertices i and j
- $Cl(i)$ = Local clustering coefficient of vertex i
- Cl_g = Global clustering coefficient
- $C_S(i)$ = Strength centrality of vertex i
- $C_C(i)$ = Closeness centrality of vertex i
- $C_B(i)$ = Betweenness centrality of vertex i
- $C_{4b}(i)$ = Bipartite clustering coefficient of vertex i
- $C^w_{4b}(i)$ = Weighted bipartite clustering coefficient of vertex i
- d_{ij} = Shortest path (geodesic distance) between vertices i and j
- $\langle d \rangle$ = Average path length across a graph
- $d_{\max}$ = Diameter of a graph

www.manaraa.com

- δ = Edge density
- $G(V,E)$ = A graph comprised of a set of $|V|$ vertices and a set of $|E|$ edges
- k_i = Degree of vertex i
- $\langle k \rangle$ = Average vertex degree
- k_{nn} = Nearest neighbours average degree
- $|E|$ = Total number of edges in a graph
- $|V|$ = Total number of vertices in a graph
- $N_{ij}^{(d)}$ = Number of paths of length d, between vertices i and j
- $\mathbb{P}(k)$ = Degree distribution
- s_{nn} = Nearest neighbours average strength
- $\mathbf{W}$ = Weighted matrix of a graph
- w_{ij} = Weight of the edge between vertices i and j

2.2 Network Basics

A *graph* (or network[1]) is a way of representing pairwise interactions among a set discrete entities [13]. Each graph $G(V,E)$ consists of a set V of N *vertices* (e.g. people, web pages, routers etc...) and the E *edges* (links) between those vertices (e.g. friendships, hyperlinks, cables). A pair of vertices are considered *adjacent* in the network if they are connected by an edge.

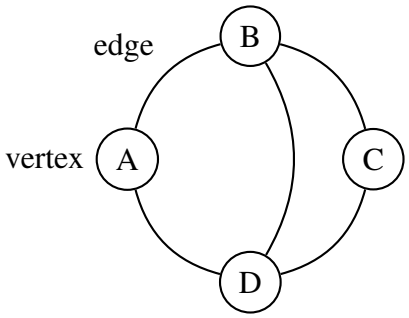
For mathematical purposes a graph is often represented through its adjacency matrix $\mathbf{A}$. Each element a_{ij} of this $n \times n$ matrix denotes whether an edge connects vertices i and j together, where

$$a_{ij} = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are adjacent} \\ 0, & \text{otherwise.} \end{cases} \tag{2.1}$$

The adjacency matrix of an undirected network is always symmetric i.e. $a_{ij} = a_{ji} \quad \forall i,j$. As such, we record two entries for each edge. For example, in Figure 2.1 the edge between vertices A and B is represented by both the elements, $a_{\mathrm{AB}} = 1$ and $a_{\mathrm{BA}} = 1$ in the adjacency matrix.

[1] Although often used as synonyms, a network typically refers to a real-world system, whereas a graph is the mathematical representation of a network [13].

www.manaraa.com



$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

Fig. 2.1 (Left) a undirected graph consisting of a set of 4 vertices {A,B,C,D} and a set of 5 edges {(A,B),(B,C),(B,D),(C,D),(D,A)}. Vertex A is adjacent with vertices B and D, but not with vertex C. (Right) the corresponding adjacency matrix **A**.

2.2.1 Vertex Degree

We often need to count the number of edges connected to a particular vertex, which we call the vertex's *degree*, k [56]. The degree k_i of vertex i can be directly obtained from the adjacency matrix. For an undirected network, each vertex's degree can be calculated by summing over either the rows or the columns of a network's adjacency matrix as follows

$$k_i = \sum_{i=1}^{N} a_{ij} = \sum_{j=1}^{N} a_{ij}, \tag{2.2}$$

where N is the total number of vertices in the graph i.e., the size of the network. A vertex degree give some indication as to its relative importance or 'centrality'. So for example, if most vertices in a graph have the same degree, we can say that each vertex plays an equal part in the network. Sometimes however, a few vertices have many more edges compared to others. These high-degree vertices are often referred to as 'hubs' [13], and are considered to play a more important or central role in the network. Degree centrality is particularly important in graphical models that predict the diffusion of information and the spread of infections [89].

The total number of edges in a graph $|E|$, can be expressed as the sum of the vertex degrees

$$|E| = \frac{1}{2} \sum_{i=1}^{N} k_i, \tag{2.3}$$

where the 1/2 factor corrects for the fact that the summation in (2.3) counts each edge twice. The maximum number of edges a graph of V vertices, is given by

$$E_{\max} = \frac{N(N-1)}{2}. \tag{2.4}$$

www.manaraa.com

A *complete* graph K_N, is one in which all possible edges exist. From these values we can also calculate a graph's *mean degree* $\langle k \rangle$, i.e., the average degree of all vertices in the network, as follows,

$$\langle k \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} k_i = \frac{2|E|}{N}, \tag{2.5}$$

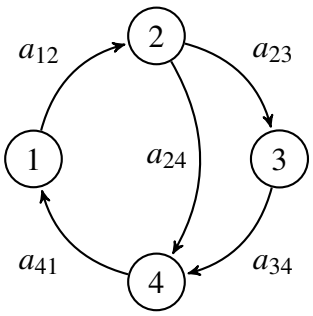
where $|E|$ is the total number of edges.

The *edge density* δ of a graph is defined by the number of edges $|E|$ divided by the maximum number of possible edges $E_{\max}$. For example, the graph in Figure 2.1 for example, has 5 edges out of a possible 12, hence the density is $\delta = 5/12 = 0.42$. The density can also be derived by dividing the mean degree by the maximum number of possible edges

$$\delta = \frac{\langle k \rangle}{N-1}. \tag{2.6}$$

2.2.2 Directed Graphs

In a directed graph, edges point from one vertex to another, and the adjacency matrix is not necessarily symmetric. That is to say. if an edge is joining vertices i to j, then there need not necessarily also be an edge from j to i, however, if vertices i and j are connected both to and from one another, two edges (i, j) and (j, i), are added to the adjacency matrix. An example of a directed graph and its associated adjacency matrix $\mathbf{A}$, is given in Figure 2.2.



$$\mathbf{A} = \begin{bmatrix} 0 & a_{12} & 0 & 0 \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \\ a_{41} & 0 & 0 & 0 \end{bmatrix}$$

Fig. 2.2 (Left) a directed graph consisting. The direction of travel is shown by an arrow on each edge. (Right) the corresponding adjacency matrix $\mathbf{A}$, where the element $a_{12} = 1$ indicates that there exist an edge from vertex 1 to vertex 2.

In a directed graph we need to distinguish between the *in-degree* k_i^{in}, representing the number of edges that point *to* vertex i, and *out-degree* k_i^{out}, representing the number of edges that point *from* the vertex i to other vertices in the network. The (total) degree is the sum of

www.manaraa.com

both in- and out-degrees

$$k_i = k_i^{\text{in}} + k_i^{\text{out}}. \tag{2.7}$$

The total number of edges in a directed network is given by

$$|E| = \sum_{i=1}^{N} k_i^{\text{in}} = \sum_{i=1}^{N} k_i^{\text{out}}. \tag{2.8}$$

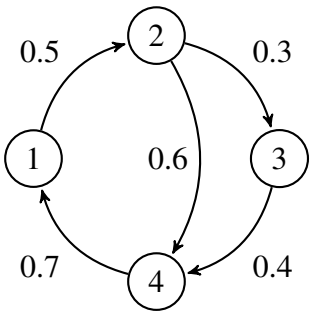
The mean degree of a directed network is simply

$$\langle k \rangle = \frac{|E|}{N}. \tag{2.9}$$

In a directed graph we also need to distinguish between the in- and out-neighbourhood of a vertex [25]. The in-neighbourhood of vertex v, denoted as $V^{\text{in}}(v)$ is the set of vertices that connect *to* v. That is, $V^{\text{in}}(v) = \{u : (u,v) \in E\}$. The out-neighbourhood of v, denoted as $V^{\text{out}}(v)$ is the set of vertices that are connected *from* v. That is, $V^{\text{out}}(v) = \{u : (v,u) \in E\}$. So for example in figure 2.2, the in-neighbourhood of vertex 2 consist only of vertex 1, whereas the out-neighbourhood contains both vertex 3 and 4.

2.2.3 Weighted Graphs

In a weighted graphs, each edge has an associated numerical value, referred to as the edge *weight*. Edge weights can be interpreted in many ways, such as a measure of the distance between two vertices [25]. The edge weights are recorded in a *weighted matrix* $\mathbf{W}$, whose elements $w_{i,j}$, denote the weight for edges (i, j) (see Figure 5.3). In most contexts, edge weights are either a non-negative real number or natural number.



$$\mathbf{W} = \begin{bmatrix} 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.3 & 0.6 \\ 0 & 0 & 0 & 0.4 \\ 0.7 & 0 & 0 & 0 \end{bmatrix}$$

Fig. 2.3 (Left) a weighted-directed graph. The weight of each interaction is shown by the value adjacent to each edge. (Right) the corresponding weighted matrix $\mathbf{W}$, where the element w_{ij} indicates the weight associated with the edge from i to j.

www.manaraa.com

2.3 Traversing Graphs

While traversing graphs is not as prominent in trade network analysis as in other network science applications, concepts such as the 'shortest path' serve as the basis for a number of important measures of network structure [25]. The concept of traversing graphs begins with noting that non-adjacent vertices may be reachable from one to the other by visiting a series of adjacent vertices.

A *path* between two vertices must follow a sequence of adjacent vertices $i \to j \to \cdots \to k$ such that each consecutive pair $i \to j$ is connected by an edge. In addition, no vertex can be visited more than once, and no edge can be repeated. The length of a path is measured by the number of edges that must be traversed to reach j from i. The shortest path between two vertices is known as the *geodesic path*, the length of which is referred to as the geodesic *distance* d. In an undirected network, the distance from i to j is always the same as the reverse path, such that $d_{ij} = d_{ji}$. However, in a directed network this is not always the case, such that $d_{ij} \neq d_{ji}$ for some pairs of vertices.

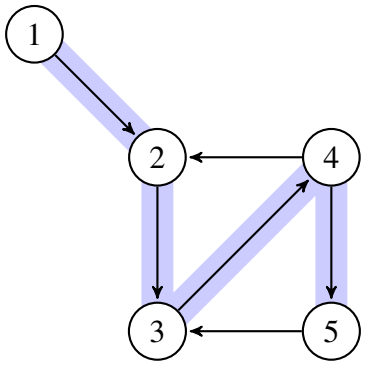


Fig. 2.4 A directed graph. The highlighted path $1 \to 2 \to 3 \to 4 \to 5$, is the geodesic distance between vertices 1 and 5.

The number of shortest paths, N_{ij}, between a pair of vertices can be determined directly from a graph's adjacency matrix $\mathbf{A}$. If there is a direct path between them (i.e., $d_{ij} = 1$), then $a_{ij} = 1$ ($a_{ij} = 0$ otherwise). For paths of length d, $a_{ik} \ldots a_{lj} = 1$ ($a_{ik} \ldots a_{lj} = 0$ otherwise). The number of paths of length d is given by

$$N_{ij}^{(d)} = \left[\mathbf{A}^d\right]_{ij}. \tag{2.10}$$

www.manaraa.com

where $\left[\ldots\right]_{ij}$ denotes the (ij)th element of $\mathbf{A}$. Equation (2.10) holds for both directed and undirected networks [13]. The distance between vertices i and j is the path with the smallest d for which $N_{ij}^{(d)} > 0$.

2.3.1 Diameter

Where a graph is fully connected, i.e., all vertices can be reached from one another along a set of paths, its *diameter* can be measured as the longest-shortest path over all pairs of vertices (n_i, n_j). For example, the graph in Figure 2.4 has a diameter of $d_{\max} = 4$. The *average path length* $\langle d \rangle$, is simply the average distance between all pairs of vertices in a graph. For a directed network of N vertices, $\langle d \rangle$ is given by

$$\langle d \rangle = \frac{1}{N \cdot (N-1)} \sum_{i \neq j}^{N} d(n_i, n_j). \tag{2.11}$$

For an undirected network, we need to multiply the right side of equation (2.11) by two.

2.3.2 Components

In many real-world networks, not every pair of vertices are necessarily connected by a path. If any two vertices are unreachable, a graph is said to be *disconnected*, and the distance is said to be either infinite or undefined. A disconnected graph is comprised of two or more *components*. Each component is defined as a *subgraph*, where each vertex has a path to all others (i.e., they are mutually reachable). The size of a component is simply the number of vertices it contains, with the one containing the most vertices referred to as the *largest component* [105]. In this case the network diameter is measured as the diameter of its largest component.

2.4 Network Measures

2.4.1 Degree Distribution

Recall that a vertex's degree is the number of edges connected to that vertex. We often want to know how many vertices in a graph have a similar degree. Making a histogram of all vertex degrees gives the *degree distribution*, which can also be defined as the fraction of vertices N, that have degree k [107]. If there are N vertices in a graph, and N_k have a degree k, the degree

www.manaraa.com

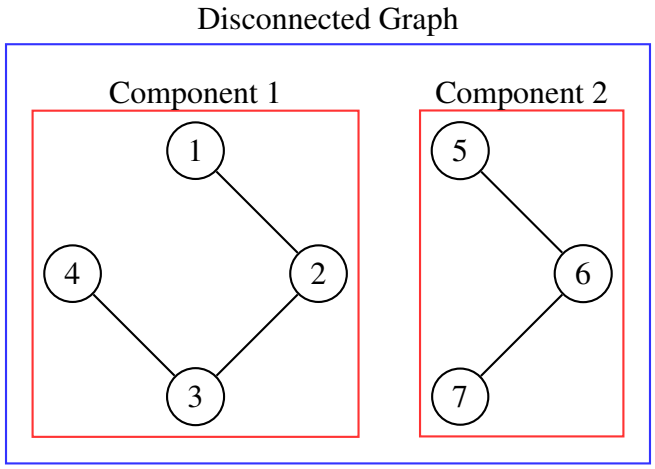


Fig. 2.5 A disconnected graph comprised of two components. The size of Component 1 is $N = 4$ whereas in Component 2, $N = 3$, thus the former is the largest component of the graph.

distribution is given by

$$\mathbb{P}(k) = \frac{N_k}{N}. \tag{2.12}$$

The degree distribution can also be thought of as the probability that a vertex chosen at random has degree k. The degree distribution has taken a central role in network theory as the precise functional form of $\mathbb{P}(k)$ is thought to influence the behaviour of many network phenomena [14], including the relative robustness of the network to attack [5]. Another reason for its importance is that the calculation of many network properties requires us to know $\mathbb{P}(k)$. For example, the average degree of a network can also be written as

$$\langle k \rangle = \sum_{k=0}^{\infty} k\mathbb{P}(k). \tag{2.13}$$

2.4.2 Degree-Degree Correlations

The in- and out-degrees of connected vertices may be correlated [106]. This relationship can be measured by the degree-degree correlation coefficient, which denotes the probability that a vertex of degree k_i is connected to a vertex of degree k_j, given by

$$r = \frac{\langle k_i k_j \rangle - \langle k \rangle^2}{\langle k^2 \rangle - \langle k \rangle^2}, \tag{2.14}$$

where averages are taken over all pairs of neighbours, i and j. This quantity takes values in the range $-1 \leq r \leq 1$, depending on the strength and direction of the relationship.

www.manaraa.com

2.4.3 Nearest Neighbours Average Degree

Another way of representing a degree-degree correlation in a graph is the *nearest neighbours average degree* given by

$$k_{nn} = \sum_{k'} k' \mathbb{P}_c(k'|k), \tag{2.15}$$

which measures the probability that an edge of a k degree vertex points to a vertex with degree k'.

2.4.4 Clustering

The *clustering coefficient* (also referred to as transitivity) is related to the number of triangles in the network (see Figure 2.6). The *local clustering coefficient* of vertex i can be expressed as

$$Cl(i) = \frac{\text{the number of closed triples (triangles) connected to vertex i}}{\text{the number of triples centred on vertex i}}, \tag{2.16}$$

where a 'triple' refers a single vertex with edges running to two other vertices. Alternatively, local clustering can be interpreted as a measure of a graph's local density, i.e., the more densely interconnected the neighbourhood of a vertex, the higher its clustering [13]. More formally, the local clustering coefficient for a undirected graph is given by

$$Cl(i) = \frac{2E_i}{k_i(k_i-1)}, \tag{2.17}$$

and for a directed graph

$$Cl_d(i) = \frac{E_i}{k_i(k_i-1)}, \tag{2.18}$$

where E_i is the number of edges between the k_i neighbours of vertex i. The local clustering coefficient is in the range $0 \leq Cl(i) \leq 1$ where, $Cl(i) = 0$ means that there are no edges between $i's$ neighbours and $Cl(i) = 1$ implies that every neighbour is connected to each other [13]. For vertices with degree 0 or 1, we set $Cl(i) = 0$.

By measuring the local clustering as a function of degree $Cl(k)$, we can reveal the presence of hierarchy across the graph [122]. For instance, if values of Cl rise with higher values of k we would see high degree vertices connected to tightly interacting neighbours and low degree vertices connected to only loosely interacting neighbours.

www.manaraa.com

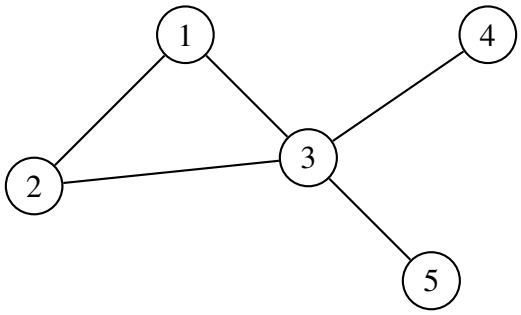


Fig. 2.6 An example graph with one triangle and eight connected triples. Each of the vertices have local clustering coefficients equal to $Cl(1) = 1$, $Cl(2) = 1$, $Cl(3) = 1/6$, $Cl(4) = 0$, and $Cl(5) = 0$. From equation (2.19), the global clustering coefficient $Cl_g = 0.4\dot{3}$.

The *global clustering coefficient* is simply the average of all local clustering coefficients across the graph

$$Cl_g = \frac{1}{N}\sum_i Cl(i). \tag{2.19}$$

This measure can also be considered the probability that two neighbours of a randomly selected vertex are themselves connected to each other [13, 35]. In random-graph models the global clustering coefficient is generally low and falls to 0 as the graph size increases. However, in many real-world networks the clustering coefficient has been found to be much higher and remains so, even in larger networks [35].

2.4.5 Centrality

Some vertices are considered to be important due to their topological position in a graph [35]. We call this measure the vertex's *centrality*. However, there are several different measures of centrality that capture different aspects of a vertex's importance. The context of the graph needs to be taken into account in order to make the right choice of centrality measure [89].

Strength Centrality

The *strength centrality* (often referred to as just vertex strength) is measured by the total weight of a vertex's edges. For vertex i, strength centrality is given by

$$C_S(i) = \sum_{j=1}^{N} a_{ij} w_{ij}, \tag{2.20}$$

www.manaraa.com

where w_{ij} is the weight of the edge between vertex i and j. The context in which this measure is applied is particularly important for interpreting the individual edge weights and vertex strength. For example, in an airport network, each edge weight would record the number of flights between two airports and the vertex strength would measure the total air traffic handled by each airport [17].

We might consider whether there is a correlation between a vertex's strength and its degree. If they are correlated, the relationship between the two can be approximated by

$$C_S(k) \sim k^{\beta}, \tag{2.21}$$

where the parameter β gives the shape of the correlation [63]. One possibility is that $\beta < 1$, which would show that those vertices with higher degrees have many edges, but with low weights. In such as graph, the importance of high degree vertices might be less than we might have otherwise expected. Edge weights could also be randomly distributed, such that $\beta \approx 1$. However, in the case where $\beta > 1$, a vertex's degree would highly underestimate its importance in the graph [63].

Average nearest-neighbour strength s_{nn} measures how intense are edge weights between the nearest-neighbours of a given vertex. Average nearest-neighbour strength of vertex i is given by

$$s_{nn}(i) = \frac{1}{N(i)} \sum_j C_S(j). \tag{2.22}$$

We can also measure a correlation between the average nearest-neighbour strength and vertex strength centrality, which gives a measure of network strength assortativity (if positive) or disassortativity (if negative) [44].

Closeness Centrality

The *closeness centrality* measures how many steps it takes, from a particular vertex, to reach any other vertex in the graph [56]. For vertex i the closeness centrality $C_C(i)$ is given by

$$C_C(i) = \sum_{j \neq i} \frac{1}{d_{ij}}, \tag{2.23}$$

where d_{ij} the distance between vertex i and j. We take the inverse of the distance in (2.23), so that larger values of $C_C(i)$ indicate more centrality, in line with other centrality measures. As

www.manaraa.com

such, a vertex with a high closeness centrality can 'communicate' or travel (depending on the context) very quickly with other vertices in the graph [89].

Betweenness Centrality

Betweenness centrality measures how often a given vertex is on a path between two other vertices [57]. The betweenness of vertex i, is defined as follows

$$C_B(i) = \sum_{j,k \neq i} \frac{d_{ijk}(i)}{d_{jk}}. \tag{2.24}$$

where d_{jk} is the number of shortest paths from j to k, and $d_{ijk}(i)$ is the number of those paths that pass through i. Dividing one by the other gives the fraction of shortest paths that pass through vertex i. For example, if we built a network representing international shipping routes, the ports (i.e., vertices) with high betweenness centrality have the ability to extort benefits from trade flowing across the network [25].

2.5 Bipartite Graphs

If the vertices in a graph represent two distinct sets of objects (e.g., buyers and sellers or importers and exporters) and only vertices in different sets can interact, then we have a *bipartite graph*. More formally, a bipartite graph G is the union of two disjoint sets U and V. If $e_{12} \in E(G)$ then $u_1 \in U$ and $v_2 \in V$. In other words, the two endpoints of any edge must be in different sets [89].

A bipartite graph can be 'projected' into two, one-mode projections, each representing one set of vertices [13]. In each projection, two vertices are connected by an edge, only if they share a neighbour in the original bipartite graph [89]. As an example we show a bipartite graph and both projections in Figure (2.7). Here, we might represent each vertex in the set U (i.e., those denoted with a number) represents a 'buyer', while each vertex in the set V (i.e., those denoted with a letter) represents a 'seller'. The edge e_{1A}, between buyer 1 and seller A therefore, could represent a sale of goods from A to 1. Note that buyers cannot purchase goods from other buyers or vice versa. Interestingly, one-mode projections economic networks can tell us which buyers or sellers are in direct competition with one another.

However, in creating one-mode projections some information is lost [89]. Consider an example where two buyers i and j both trade with the same two sellers (not shown). Here, a

www.manaraa.com

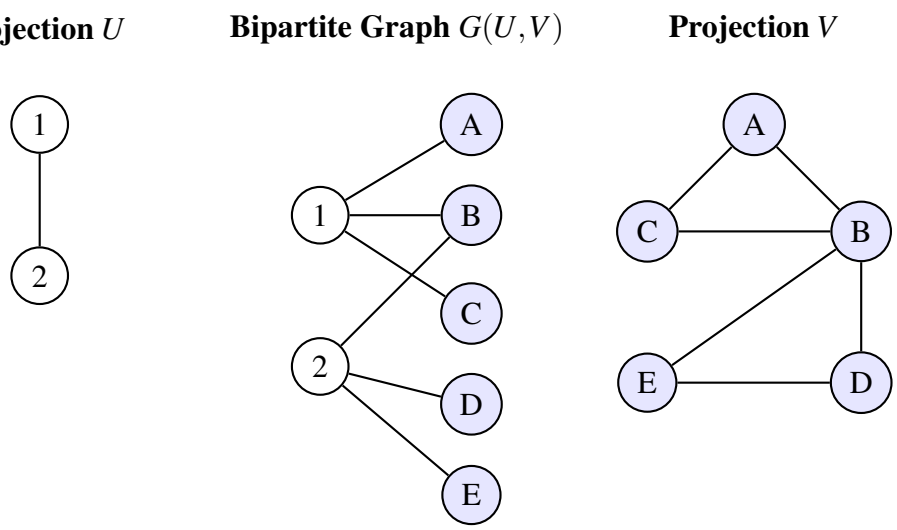


Fig. 2.7 (*Centre*) a bipartite network comprised of two sets of vertices, U and V. Vertices in U connect only to vertices in V. Hence there are no direct U–U or V–V edges. (*Left and right*) one-mode projections of U and V. Projection U is obtained by connecting two U-vertices to each other if they share an edge to the same V-vertex and projection V is obtained by connecting two V-vertices to each other if they share an edge to the same U-vertex.

one-mode projection of the buyer set would show only one edge between i and j, where there is in fact two paths between them in the bipartite graph. As such, we need to take into account the number of paths (of length 2) connecting each set of buyer vertices in the bipartite graph. If we define **B** as the adjacency matrix of the buyer projection, then number of paths can be assigned as the weight w_{ij} of the edge between i and j. Each element of the buyer's adjacency matrix **B** therefore becomes

$$b_{ij} = \begin{cases} w_{ij}, & \text{if } i \text{ and } j \text{ are connected with weight } w_{ij} \\ 0, & \text{if } i \text{ and } j \text{ are not connected.} \end{cases} \tag{2.25}$$

A similar adjacency matrix can also be derived for the seller's projection.

2.5.1 Bipartite Clustering

By the standard measure, the clustering coefficient in bipartite graphs is always zero. To get around this problem, clustering is often calculated on a one-mode projection of a bipartite graph. However, as previously noted, using a one-mode projection leads to a loss of information and, in some cases, can result in changing the meaning of what edges in the network are supposed to represent [63]. An alternative option proposed by Lind *et al.* [92], is to calculate a *bipartite*

www.manaraa.com

clustering coefficient using cycles of four connections (i.e., quadrilaterals rather than triangles). In this case, we call all vertices connected to i, first neighbours, and all vertices connected to i's first neighbours, the second neighbours (even though they are not directly connected to i). Using this schema, the bipartite clustering coefficient can be defined as the fraction of common edges between second-nearest neighbours (not counting i), and calculated as follows

$$C_{4\mathrm{b}}(i) = \frac{q_i}{k_i^{\mathrm{nn}} \times k_i(k_i-1)/2}, \tag{2.26}$$

where q_i is the number of observed quadrilaterals around i, k_i^{nn} is number of second-nearest neighbours, and $k_i(k_i-1)/2$ total number of pairs of neighbours (see Figure 2.8).

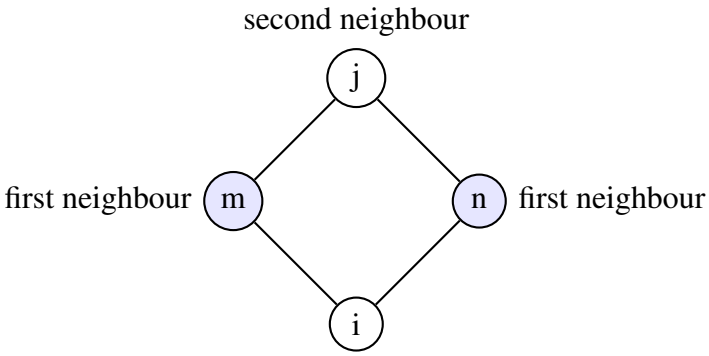


Fig. 2.8 Graph showing a quadrilateral q_i formed by vertex i and vertex i's first nearest neighbours (m,n) and second nearest neighbour (j).

In addition to finding the number of clusters in a graph, we might also want to know whether those clusters are formed by strong or weak edges. For this we need to take into account the weight of the edges that form each quadrilateral. For this purpose, Gilarranz et al. [63] developed a weighted clustering coefficient, $C_{4\mathrm{b}}^{\mathrm{w}}(i)$, which is a generalisation of $C_{4\mathrm{b}}(i)$. In this weighted case, the number of quadrilaterals is multiplied by the average normalised weight[1] of the two edges connecting vertex i as follows

$$C_{4\mathrm{b}}^{\mathrm{w}}(i) = \frac{\sum_{m,n} q_{imn}\left(\frac{\tilde{w}_{im}+\tilde{w}_{in}}{2}\right)}{k_i^{\mathrm{nn}} \times k_i(k_i-1)/2}, \tag{2.27}$$

where q_{imn} is the number of quadrilaterals formed with i's nearest neighbours m and n, and $\tilde{w}_{im}$ and $\tilde{w}_{in}$ are the normalised weights of the edges between i and m and between i and n. Normalisation can be calculated by dividing each edge weight by vertex i's degree over the

[1] Normalising the weights ensures that the clustering coefficient will take a value between 0 and 1.

www.manaraa.com

sum of its edge weights (i.e., the vertex's strength centrality) as follows

$$\tilde{w}_{im} = \frac{w_{im}}{\left(\frac{C_S(i)}{k_i}\right)}. \tag{2.28}$$

2.6 Generative Network Models

Studies of empirical networks have found that despite differences in the nature of their vertices, many networks exhibit similar topological properties. This is important, as a common architecture may point to a common pattern of network formation [18]. Two generative models often seen in the trade network literature are the Erdős and Rényi (E-R) *random graph* model [41] and the Barabási-Albert (B-A) *preferential attachment* model [14].

2.6.1 Random Graph Model

In the E-R random graph model, each pair of vertices is connected by an edge, with probability p [41]. The degree distribution of a random network follows the binomial distribution

$$\mathbb{P}(k) = \binom{N-1}{k} p_k (1-p_k)^{N-1-k}, \tag{2.29}$$

where p_k is the probability that a vertex has k edges, and $(1-p)^{N-1-k}$ is the probability that the remaining $(N-1-k)$ edges are missing. Where networks are sparse, the degree distribution can be approximated by a Poisson distribution as follows

$$\mathbb{P}(k) \sim e^{-k} \frac{\langle k \rangle^k}{k!}. \tag{2.30}$$

As such, the connectivity of vertices in this model is very homogeneous and the tail of this distribution is narrow [18]. We can also say that the probability of a vertex having a degree larger than the mean $\langle k \rangle$, decays exponentially. The E-R model is useful in that the expected structure of graph can be specified through p. For example, where $p = 0$, there are no edges and the graph empty. Alternatively, where $p = 1$, all possible edges are connected making complete graph. However, we can also move incrementally, from $p = 0$ to $p = 1$ and measure how various network statistics change along with p. Real-world applications of the E-R model are limited however, as the majority of empirical networks display heterogeneous degree distributions [18].

www.manaraa.com

2.6.2 Preferential Attachment Model

The B-A preferential attachment model is thought to better reproduce the degree distributions of many real-world networks, which suggests that real networks might be generated by similar processes [32]. The preferential attachment works by attaching a new vertex to an existing vertex, with a probability proportional to the existing vertex's degree [18]. This model has been usedacross many domains, including biological, physical and social networks [13]. In graphs generated using this model, vertices in the tail of the distribution are far more connected than would be expected by chance [107]. In this model, the network's degree distribution $\mathbb{P}(k)$ is approximated by a power-law

$$\mathbb{P}(k) \sim k^{-\gamma}, \tag{2.31}$$

for values $\gamma > 0$, where γ is the scaling parameter of the distribution. If we take a logarithm of Eq. (2.31), we obtain

$$\log\mathbb{P}(k) \sim -\gamma \log k + c, \tag{2.32}$$

where c is a constant. Therefore, $\log \mathbb{P}(k)$ is expected to depend linearly on $\log k$, with the slope of this line given by γ, as observed in Figure 2.9. This kind of network is often referred to as having a *scale-free* structure. The term 'scale-free' is used here, as the relationship between degree and degree distribution is not defined on any particular scale [18].

In Figure 2.9 we show the respective degree distributions for two simulated networks, the first using the E-R model [41], and the other created using the B-A model [14]. In both cases the resulting networks have $N = 1000$ vertices, and approximately the same mean degree $\langle k \rangle \approx 8$, i.e., a typical vertex in either network is connected by eight edges. Given that values of $\mathbb{P}(k)$ and k differ by orders of magnitude, we use a *loglog* plot, in which we show $\log(\mathbb{P}(k))$ as a function of $\log(k)$. We can see that despite having the same values for N and $\langle k \rangle$, the two networks have widely different degree distributions. For the random network, vertex degrees are distributed equally around the mean, with a maximum degree of only $k_{\max} = 18$. As per equation (2.30), we observe the exponential decay in the tail of the distribution. The range of values of k are much broader across the B-A model network. One vertex in particular has almost 500 edges, connecting it to half the network. However, these high degree vertices are very rare in scale-free networks, with the vast majority of vertices having only a few edges [63]. In our example, 85% of vertices have $k \leq 8$ (i.e., fewer than the mean degree $\langle k \rangle$).

www.manaraa.com

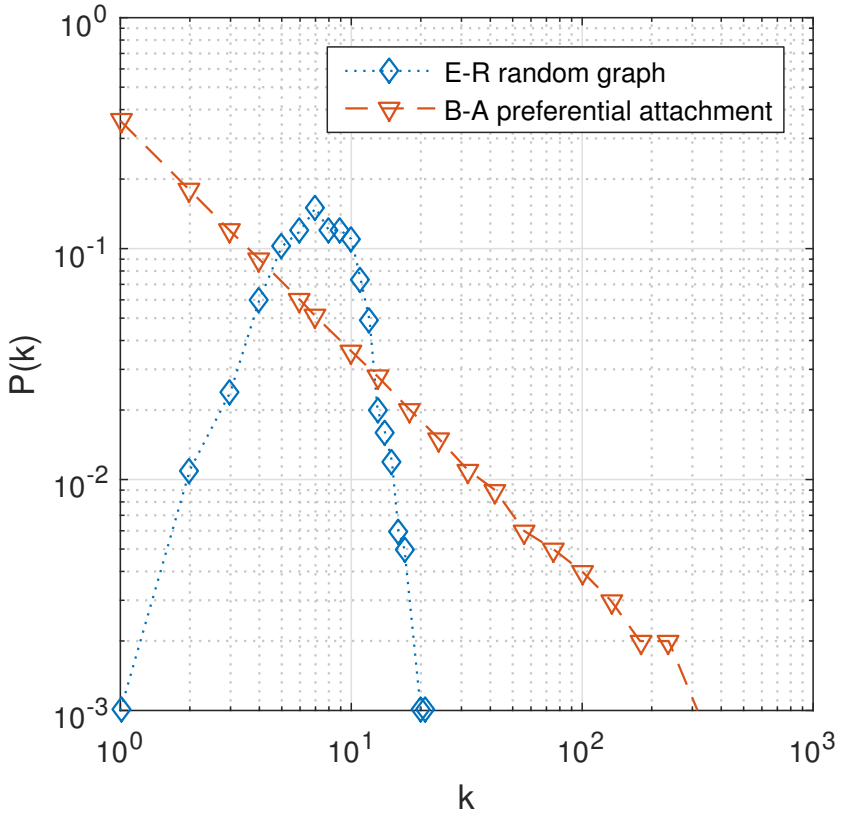


Fig. 2.9 The respective degree distributions of two graphs simulated using the E-R model [41] and the B-A model [14].

www.manaraa.com

Chapter 3

Review of the Trade Network Literature

3.1 Introduction

Broadly speaking, the literature on the network structure of trade follows a singular approach, as the objective of most authors is the same: to develop an appropriate network representation of the international trading system, and explore the network's statistical properties by applying this model to empirical data. The result is a great number of different network models and statistical measures that would be hard to summarise here without going to excessive length. Rather than listing all the different models in the literature, we begin by outlining a generalised trade network model that incorporates the most common elements from across the field of economics. This model can then serve as a benchmark from which we can compare how and where authors differ in their approach. In Section 3.2, we outline the various caveats to this generalised model, and detail how network models have been developed to account for binary (3.2.1), weighted (3.2.2), and bipartite (3.2.3) interpretations of international trade. In Section 3.3 we explore how the implementation of these models affects the topology of trade through reported statistical properties, including the degree distribution (3.3.1), degree-degree correlations (3.3.2), clustering coefficient (3.3.3), and strength centrality (3.3.4).

The generalised trade network model can be described as follows. Consider a market of n countries trading goods among themselves. On a graph, each country can be represented by a vertex and trade between each pair of countries can be represented by an edge. If country i exports goods to country j, then a directed edge is drawn from i to j, so that the element $a_{ij} = 1$ in the corresponding $n \times n$ adjacency matrix $\mathbf{A}$. Otherwise, no edge is drawn and $a_{ij} = 0$ [62]. As a result, rows of the adjacency matrix represent countries that export goods, and columns represent countries that import goods [43]. In addition, the sum of each row in the adjacency

www.manaraa.com

matrix gives the out-degree $\sum_j a_{ij} = k_{\text{out}}$, and the sum of each column gives the in-degree $\sum_i a_{ij} = k_{\text{in}}$. If the trade data spans more than one time period, multiple graphs can be created, where each adjacency matrix $\mathbf{A}(t)$ represents a snapshot of the network, and each element $a_{ij}(t)$ denotes whether goods where traded between i and j at time t. This time dependent model is often used, as the structure of the international trade network continuously changes due to variation in the number of countries trading and due to the continuous rearrangement (or 'rewiring' in network terminology) of bilateral trade relations [62].

3.2 Network Construction

The varying network interpretations of international trade can be summarised as follows:

3.2.1 Binary Trade Networks

One of the earliest works on modelling international trade as a network, is that of Serrano and Boguñá [122]. Here, trade is studied as a *binary* (i.e., unweighted) network. This is to say the authors do not consider the 'edge weights' of the network in their analysis which can be used to represent either the volume or value of trade between each pair of countries [43]. A binary network can be recorded as either *undirected* or *directed*, depending on whether the direction of trade is important to the analysis. In the undirected case, any trade, from either country i to j, or from country j to i, is represented as an edge, meaning that in the adjacency matrix $\mathbf{A}$ of the graph is symmetric $a_{ij} = a_{ji}$, and $a_{ii} = 0, \quad \forall i$. Alternatively, in the directed binary case, the direction of trade is specified in the adjacency matrix, such that elements $a_{ij} = 1$ if trade is directed from vertex i to vertex j, and $a_{ij} = 0$ otherwise [128].

To create this network model, Serrano and Boguñá [122] extracted data from the UN Comtrade database [141], which has become the primary source of trade data across the literature. These data contain, for each country and each year, a list of imports and exports. Each entry details the type of good traded, the corresponding source/destination, the volume (in Millon Metric Tonnes, MMT) and value (in thousands of US$) of the goods, and the UN Harmonized System (HS) classification code that identifies the class of goods traded (e.g., agricultural, electronics, chemicals, etc.). In this specific case, the top forty most exchanged goods in the year 2000 were selected as a sample of the international trade network. In the Comtrade data, an asymmetry is noted between the lists of imports and exports, so it is common to find entries listing exports of a good from country i to country j, but not the corresponding

www.manaraa.com

entry of imports in country j, from county i [122]. These errors are due to the fact that there is no global reporting agency for trade data, rather the onus is on exporting countries to report exports and importing countries to report imports. Given the voluntary nature of this reporting and the vast number of transactions taking place every year, these records are unlikely to be 100% accurate. Many researchers in the literature simply chose one list over the other, Squartini *et al.* [128] and Fagiolo *et al.* [45] use the import list for example. However, the export data is considered more accurate than imports, based on the UN statistics division's own analysis [141], which we use throughout this thesis. Serrano and Boguñá [122] take a different approach. They construct two alternative adjacency matrices from the data, $\mathbf{A}^{\text{imp}}$ using only the list of import transactions, and $\mathbf{A}^{\text{exp}}$ using only the export transactions list. Each element in these adjacency matrix is defined so that $a_{ij}^{\text{imp}} = 1$ if country i imports from country j, and $a_{ji}^{\text{exp}} = 1$ if country j exports to country i. Thus where the data are in agreement, the elements $a_{ij}^{\text{imp}} = a_{ji}^{\text{exp}}$. They combine these adjacency matrices to create a new adjacency matrix $\hat{\mathbf{A}}$, that completes the symmetry relation between imports and exports. Mathematically, this method uses the Kronecker delta function, which is typically used to denote where two variables are equal, and can be written as,

$$\delta_{i,j} = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases} \tag{3.1}$$

Using this function the complete adjacency matrix can be calculated as follows,

$$\hat{\mathbf{A}} = \frac{1}{1 + \delta_{(\mathbf{A}^{\text{imp}} + \mathbf{A}^{\text{exp}}),2}} [\mathbf{A}^{\text{imp}} + \mathbf{A}^{\text{exp}}]. \tag{3.2}$$

Here, the Kronecker delta function $\delta_{(\mathbf{A}^{\text{imp}} + \mathbf{A}^{\text{exp}}),2}$ will return 1 where the adjacency matrices are in agreement and 0 otherwise. Thus where $a_{ij}^{\text{imp}} = a_{ji}^{\text{exp}}$, the first part of equation (3.2) will be $\frac{1}{1+1} = 0.5$, and the later $a_{ij}^{\text{imp}} + a_{ji}^{\text{exp}} = 2$. This gives each symmetrical element in the complete adjacency matrix as $\hat{a}_{ij} = 1$. In addition, where only one country reporting an import/export transaction, the corresponding element in the complete adjacency matrix will also equal 1, given that the first part of equation (3.2) will be $\frac{1}{1+0} = 1$, and the later $a_{ij}^{\text{imp}} + a_{ji}^{\text{exp}} = 1$. Only where $a_{ij}^{\text{imp}} + a_{ji}^{\text{exp}} = 0$, will the complete adjacency matrix show no trade relationship. In this way the complete adjacency matrix shows all connections that are relevant to at least one of the two countries involved in each transaction [122]. Where neither country has recorded a transaction, it is assumed that no trade took place. In the formulation above, edges are only drawn where the volume of trade is strictly positive [128], so that each element of the network's

www.manaraa.com

adjacency matrix is given by,

$$a_{ij} = \begin{cases} 1 & \text{if } w_{ij} > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

where the edge weight w_{ij} is typically given in terms of total US$ worth of trade. However, many others (see [153, 63] for examples) set a stricter threshold, when converting the weighted data to a binary network. This threshold θ, is used to determine where there is a *significant* amount of trade between pairs of countries [122], so that equation (3.3) becomes,

$$a_{ij} = \begin{cases} 1 & \text{if } w_{ij} > \theta, \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

Determining the value at which a given edge should be considered 'significant', is both a subjective and contentious subject. Filtering edges based on either the value or volume of trade is always going to remove some countries from the network. Doing so has a knock-on effect in terms of censoring the resulting degree k of small or poor countries. This goes against one of principal arguments in favour of the network approach, in that, every country should be considered important, regardless of its size or wealth [122]. Unfortunately, there is no method for deriving a threshold value that is not somewhat arbitrary, which is one of the main limitations of the binary network analysis.

In the model above, trade is aggregated, so that all goods traded between two countries are measured by a single value. However, Squartini *et al.* [128] decompose the global trade data into the exchange of many individual commodities. So that, while vertices still represent countries, the edges now represent the import/export of only a single commodity: such as wheat, crude oil, or aluminium for example. Using this disaggregated data, Squartini *et al.* [128] developed a new model with 97 different goods. In addition, the change in network topology is measured over time, in a similar fashion to the example in the introduction of this chapter. So for year t, the value (in US$) of good g traded from country j to country i is given by $w_{ij}^{g}(t)$. In aggregate, the total value of all goods traded is calculated as,

$$w_{ij} \equiv \sum_{g=1}^{G} w_{ij}^{g}(t) \tag{3.5}$$

www.manaraa.com

For the undirected-binary model, the value of goods traded is reinterpreted as a combination of both imports and exports as follows,

$$w_{ij} \equiv \left[\frac{w_{ij}(t) + w_{ji}(t)}{2} \right]. \tag{3.6}$$

3.2.2 Weighted Trade Networks

The main critique levelled at binary trade network models, is that (after filtering) it treats all edges as if they were completely homogeneous [43]. Across many of the networks studied (see [43, 123, 63, 21, 51]), vast differences (often by several orders of magnitude) are reported in the volume (or value) of trade between different pairs of vertices. Thus it is clear that some trade relationships are more important than others to the system as a whole. In order to take into account the existing heterogeneity in the intensity of trade, weighted-network analyses have been developed [17]. In Fagilo *et al*. [43] for example, each edge is assigned a 'weight' $w_{ij} > 0$ in a weight matrix $\mathbf{W}$, proportional to the amount of trade flowing through it. A weighted network model is therefore, fully described by a set of $n \times n$ adjacency and weight matrices $\{\mathbf{A}, \mathbf{W}\}$. In the directed case, these weightings are typically measured in US$ values or in metric tonnes. Alternatively, the value of trade can be rescaled relative to the total trade flow (see [128]) to give a percentage, as follows

$$\hat{w}_{ij} = \frac{w_{ij}}{\sum_i \sum_{j \neq i} w_{ij}}, \tag{3.7}$$

while in the undirected case

$$\hat{w}_{ij} = \frac{w_{ij}}{\sum_i \sum_{j<i} w_{ij}}. \tag{3.8}$$

In another approach, Serrano *et al*. [123] calculate the *total trade flux* across an edge. Here, the flux f_{ij} is given by the difference between the quantity of goods imported and exported by i to and from j, with the direction of the edge pointed towards the net flow. The flux is weighted by the total net trade balance, given by

$$\hat{w}_{ij}^{\text{flux}} = w_{ji} - w_{ij}. \tag{3.9}$$

A corresponding flow matrix $\mathbf{F}$ for the weighted, directed network is constructed such that each element $f_{ij} = \hat{w}_{ij}^{\text{flux}}$, if $\hat{w}_{ij}^{\text{flux}} < 0$, otherwise $f_{ij} = 0$ if $\hat{w}_{ij}^{\text{flux}} \geq 0$, i.e., the direction of each edge is towards the country with the positive net trade balance. An alternative interpretation of this method is presented in Duenas and Fagiolo [40], where the authors also look at the *absolute*

www.manaraa.com

relative trade balance, defined as

$$\hat{w}_{ij}^{\text{rel}} = \frac{|w_{ji} - w_{ij}|}{w_{ji} + w_{ij}}, \tag{3.10}$$

where the bars around the numerator $|w_{ji} - w_{ij}|$, signify to take the absolute value. While the net trade balance in equation (3.9) is measured in US$ of goods, the absolute relative trade balance in equation (3.10) has no units and is on a scale $[0,1]$ [40]. While both are measures of bilateral trade balances, they have different economic interpretations. The net trade balance simply accounts for the balance of trade in U.S. dollars, whereas, the relative trade balance accounts for the asymmetry of trade independently of the traded volume. Thus , it is a measure of *reciprocity* in bilateral trade relationships, such that, the extreme values, zero or one, are signals of strong symmetry or asymmetry in bilateral trade, respectively [40]. Note that here, a large net trade balance does not necessarily imply a high relative trade balance. For example, where $\hat{w}_{ij}^{\text{rel}} = 1$, we can say that trade is flowing in one direction only. However this says nothing about the amount of goods being traded [40].

As previously noted, there are discrepancies in the Comtrade data between the lists of imports and exports [122]. Sometimes this leads to four different quantities being reported: the value of trade exported from i to j as reported by county i which gives w_{ij}^{exp}, exports from j to i as reported by county j giving w_{ji}^{exp}, and the same with imports reported by i giving w_{ij}^{imp}, and reported by j giving w_{ji}^{imp}. Note that this situation only occurs when considering aggregate trade, as exporters of specific commodities very rarely trade with one another [51]. Bhattacharya *et al*. [21] deal with this problem in the weighted-undirected case by defining the edge weight (where weight here is a measure of the total trade volume between the two countries in US$) as follows,

$$\hat{w}_{ij} = \frac{w_{ij}^{\text{exp}} + w_{ji}^{\text{exp}} + w_{ij}^{\text{imp}} + w_{ji}^{\text{imp}}}{2}. \tag{3.11}$$

This procedure tends to average out the aforementioned discrepancies in the quantities reported [21]. Fagiolo *et al*. [43], define weightings somewhat differently, in that the value of trade across an edge is first divided by the exporting country's GDP, to give the importance of an edge relative to the exporting country's economic size [43]. This is an interesting choice, as a country's GDP is not typically considered a 'network' measure. Conventionally, GDP is defined as "the total market value of all final goods and services produced in a country in a given period, equal to total consumer, investment, and government spending, together with

www.manaraa.com

the net value of foreign trade" [99], where the net value of foreign trade is simply the value of goods exported subtracted by the value goods imported. This can be simplified into two main terms that contribute to a county's GDP: an endogenous term, often referred to as *internal demand* which is determined by the production and spending due to a country's domestic economic processes, and an exogenous term—*net trade*, determined by interactions with other countries [60]. The latter of these terms has been reinterpreted using network terminology, which allows for a partial connection to be drawn between GDP and a country's position in the trade network.

Continuing from the definition offered by Fagilo *et al.* [43], where a weighted network is described by a set of $n \times n$ adjacency and weight matrices $\{\mathbf{A}, \mathbf{W}\}$, with each element of the weight matrix proportional to the US$ amount of trade flowing through it. A proportion of a country's GDP can be determined quantitatively through the following equations [60]. First, the total import and export trade values can each be measured as a country's strength centrality $C_S(i)$ (see Chapter II), and calculated by summing over the rows and columns of the weight matrix $\mathbf{W}$, so that total imports is given by

$$C_S^{\text{imp}}(i) = \sum_{i=1}^{N} a_{ij} w_{ij}, \tag{3.12}$$

and total exports, by

$$C_S^{\text{exp}}(i) = \sum_{j=1}^{N} a_{ij} w_{ij}, \tag{3.13}$$

where w_{ij} is the weight (measured in current $US of trade) assigned to the edge between countries i and j. The *net trade* (or 'net strength') for country i is therefore given by

$$C_S^{\text{net}}(i) = C_S^{\text{imp}}(i) - C_S^{\text{exp}}(i). \tag{3.14}$$

To reach the final calculation of GDP, the net trade for country i, as follows

$$\text{GDP}(i) = \text{internal demand}(i) + C_S^{\text{net}}(i), \tag{3.15}$$

where *internal demand* values are typically taken from datasets of public accounts. We can see that the above definition anticipates that the GDP is affected by the network structure of trade. To what extent, depends on the share of GDP accounted for by net trade for each country, but also the number of countries connected together via trade across the network (measured

www.manaraa.com

by each country's degree k, or globally by the network's edge density δ). We know that the share of GDP accounted for by net trade has been increasing in recent decades. Using data dating back to 1870, Maddison [95] found that, historically, the average balance between net trade and internal demand fluctuated around a constant value, even while global GDP was increasing. This changed however, in the 1960s, when a transition towards 'globalisation' took place, which saw the rapid integration of many developing economies into the global trading system. For instance, the share of international trade that accounted for China's GDP rose from 5% in 1970, to 65% in 2006 [148]. From the 1960s onward, Maddison [95] record average net trade figures growing proportionately with world GDP, meaning that international trade has become the dominant factor in calculating changes in GDP.

However, this change in GDP is not necessarily uniform across all countries. Several country specific factors need to be taken into account. For example, international trade tends to be more important for countries that are small geographically, or by population [108]. In contrast, countries that are large and self-sufficient tend not to rely on trade and therefore see internal demand as the primary factor in GDP calculations [108]. The trade network itself can also be affected by geopolitical factors. Being surrounded by neighbouring countries with open trade regimes for instance, can also increase the share of net trade in a country's GDP [108]. There are also several possible interrelated feedback dynamics between net trade and internal demand that could effect the calculation of GDP. For instance, imports can be used as the inputs to domestically produced goods, which in turn can lead to increased internal demand [60]. Imports can also free up resources that would have been allocated to producing goods in which a country did not have a comparative advantage[1]. As such, these resources can be reallocated to more productive domestic industries, and again, increasing internal demand [60]. Trade can also attract financial institutions to invest between countries. This Foreign Direct Investment (FDI) can indirectly lead to increases in domestic business investments, government spending, and private consumption [126].

These linkages led Garlaschelli and Loffredo [61] to test whether the GDP of each country[2] can be used as a 'hidden' *fitness* variable which determines a country's position in the network (measured here by the country's degree k), as well as correlations between second-order

[1] Comparative advantage was proposed by Ricardo *et al*. [115], and states that countries have natural advantages, called 'factor endowments' such as, land area, minerals, population, and so on, that help them produce certain goods better than others.

[2] The authors use total GDP rather than the per capita GDP (i.e., divided by the total population) which is more commonly used in economics. Total GDP is a more accurate indicator of the aggregate economic activity of a country [126].

www.manaraa.com

topological properties, such as the average nearest neighbour degree $k_{nn}(k)$, local clustering coefficient $Cl(k)$. The hidden fitness model assumes that the probability $\mathbb{P}(a_{ij})$ of an edge from i to j is a function of the values of a fitness variable ω assigned to each vertex and drawn from a given distribution [61]. The fitness variable ω of country i, is calculated by normalising i's GDP by the total GDP of all countries in the network, such that $\omega(i) = GDP(i)/\sum_j GDP(j)$ [61]. Note that in the model, a free parameter β is also introduced, such that the probability of an edge from i to j is fully given by

$$\mathbb{P}(a_{ij}) = \frac{\beta\omega_i\omega_j}{1+\beta\omega_i\omega_j}, \tag{3.16}$$

and where β is tuned so that the model reproduces the same number of edges observed in the trade network [61]. Note that this model, is in effect, a binary networked version of the well known Gravity model of trade [135], where the sum of GDP values of each pair of countries is considered the most important determining factor that predicts the *volume* of trade recorded in the Comtrade data.. Using a fitness model, Garlaschelli and Loffredo [61] find that the predicted edge probability function matches well with empirical data, such that the degree of country k_i was found to be an increasing function of ω_i and thus ultimately, GDP(i). This is to say, richer countries tend to have more trade connections than poorer ones. Serrano and Boguñá [122] also measured this link between degree and GDP, and found the correlation to be relativity high, at 0.65, noting that many least developed countries (LDCs) such as, Angola, Somalia, Rwanda, and Cambodia all had low degrees, and that developed industrial countries were typically highly connected. However, they did note a significant number of cases in the reversal situation, that is, low per capita GDP countries with a large number of connections, with Brazil, China, and Russia being exemplars, in that they are rapidly developing, export orientated economies with relatively low GDP per capita, but very high degrees [122]. Similar general correlations are found for both the average nearest neighbour degree and the clustering coefficient [61]. This shows that GDP, and the distribution of GDP across the network may go some way to determining many topological properties of international trade. It is likely that countries with higher GDP have more ability to foster trade due to the linkages outlined above. However, it should be noted that the opposite is also likely true in that a country's links to international trade may also stimulate an increase in internal demand (cheaper imports leading to increased productivity for example) [60].

www.manaraa.com

3.2.3 Bipartite Trade Networks

Fernandez *et al*. [51, 52], propose a bipartite approach to modelling international trade drawing an analogy to mutualistic[1] networks (see [63] for the ecological perspective on networks), which are bipartite by nature. Here two different species (plants/pollinators for example) compete among themselves for the opportunity to cooperate across species. Similarly in trade networks, exporting countries compete among themselves for the opportunity to trade with importing countries, and vice versa. Fernandez *et al*. [51] propose that self-interest may be a common root cause of mutualism in both systems. In ecological systems, species are driven by the need to survive the process of natural selection and perpetuate the growth of its own species, and mutualistic strategies appear as a natural response to these pressures. In commodity markets, on the other hand, businesses are driven by profit maximization in order to ensure economic survival and growth. One way that businesses can achieve these aims, is to reach out across international borders to find new markets or suppliers [51]. From these observations, the authors hypothesise that exporters will be connected in the network based on a different decision-making process than that of importers, and consequently, importers and exporters will play different roles in the network's topology.

It is important to note that this theory only applies when considering individual commodity markets, and not international trade in aggregate. In the latter case, the adjacency matrices of aggregated trade networks were found to be almost entirely symmetric, with approximately 93% of directed edges being reciprocated [44]. So that, if country i exports to country j then country j almost always exports to country i. This means that, at this level, there is no significant difference between importing and exporting countries. However, the opposite is found at the individual commodity market level, where importers almost never export the same good, or vice versa [51]. Thus if country i exports a specific good to country j, then country j will likely not export that same good to country i. These findings can be explained by considering Richardo's theory of comparative advantage [115]. In his 1817 treatise, "On the Principles of Political Economy and Taxation", Ricardo posited that countries in an international trading system gain the most benefit from specialising in exporting only those goods in which they have a comparative advantage, and importing those goods in which they do not. A comparative advantage is derived from a set of so-called 'factor endowments', land, minerals, population, and so on, which give a country the ability to produce certain goods at a lower cost that their competitors. In this way, countries can pay for an array of imported goods by exporting highly

[1] Mutualism is a cooperative relationship (symbiosis), beneficial to both species involved (as opposed to parasitism) [99].

www.manaraa.com

specialised goods at competitive prices. Thus, at the individual commodity level, it is correct to assume exporters trade based on a different decision-making process than that of importers. Whether or not this has an effect on the network's topology can only be answered empirically.

Fernandez *et al.* [51, 52], do not explicitly define a bipartite trade network model, but following Gilarranz *et al.* [63], upon which their work is based, it can be formulated as follows. A bipartite graph, $G = (U, V, E)$, is drawn, comprising two sets of vertices, U consisting of $\{1, 2, \ldots, j\}$ exporters, and V consisting of $\{1, 2, \ldots, i\}$ importers, as well as a set of edges E connecting importers to exporters. Each element $a_{ij} = 1$ in the graph's adjacency matrix $\mathbf{A}$, where importer i trades with exporter j, otherwise $a_{ij} = 0$. In a bipartite system, trade is not possible with others in the same set (i.e., importers are only allowed to trade with exporters), so that all permitted edges in E are between sets U and V. Naturally, a bipartite model requires adjusted network measures (see Chapter II for detailed descriptions of the calculations required) to take into account the fundamental differences in topology. The results of using this method are reported alongside the standard network measures in the sections below.

3.3 Reported Statistical Properties

The reported statistical properties of empirical trade networks can be summarised as follows:

3.3.1 Degree Distribution

A ubiquitous feature of literature are reports of extremely high levels of heterogeneity in the degree distribution of the networks under study [19, 43, 54, 128, 129, 122, 21, 45]. The degree distribution $\mathbb{P}(k)$, measures the probability of a randomly chosen vertex to have k connections to other vertices [122]. In both the directed and bipartite cases, the degree distribution is taken for both the in-degree, representing imports and the out-degree, representing exports [43]. The resulting distributions are typically approximated by a power-law, such that $\mathbb{P}(k) \sim k^{-\gamma}$, where γ is the scaling coefficient of the distribution (see Chapter II for a technical description of these concepts). On *log-log* axis, the parameter γ also gives the slope of a linear model fit. Few authors give estimates for this parameter, owing to the difficulties of fitting power-law models to such small n observations, but those that do, find that in aggregate trade networks $\gamma \approx 2.6$ [122], whereas in individual commodity markets the exponent is typically found in the range $\gamma \approx 1.3 \pm 0.2$ [19]. In real terms, this means that the majority of countries in the international trade network have very few trade partners, whereas, countries in the tail of the

www.manaraa.com

distribution are far more connected than would be expected by chance [44]. Naturally, the causes and consequences of this large inequality in trading opportunities has become an issue of interest to economists and international development researchers alike.

Baskaran et al. [19] point to the Heckscher-Ohlin (H-O) model [110] of international trade as a possible explanation for these empirical results. The H-O model is, in essence, a formalised version of Ricardo's theory of comparative advantage. The central theorem of the H-O model states that bilateral trade flows between two countries are due to differences in what Ricardo calls 'factor endowments' (see above). The theorem goes on to state that a country will export those goods that use factors which are in abundant supply, and import goods that use factors that are scarce [100]. Baskaran et al. [19] argue that, given the H-O theorem, only a few countries will specialise in exporting a specific good whereas many countries will import those goods for which they lack the required factor endowments to produce. Therefore in a network setting we should expect each market to be comprised of a small group of highly specialised exporters connected to a larger group of non-specialised importers. While this theory does suggest a high level of degree heterogeneity is to be expected, it is not clear why a power-law distribution in particular, should be observed.

Observations of power-law degree distributions also imply a strong deviation from the classical random graph theory of Erdős and Rényi [41] (see Chapter II), which would predict an exponential decay [122]. The empirical results rather suggest that the 'preferential attachment' model of network growth is a more likely explanation [13] (see Chapter II), as the exponent γ is found to be within the range defined by many other scale-free networks [122, 19, 13]. Serrano and Boguñá [122] hypothesise that the preferential nature of growth in these trade networks is likely to be due to risk-averse consumers preferring to purchase goods from well-established exporters. However, they also anticipate that the preferential attachment mechanism could differ between countries, depending on the particular political and economic situation [122].

One issue that has gone largely unremarked in the literature is the number of deviations from the power-law behaviour of trade network degree distributions. For example, Serrano and Boguñá [122] observe that in their data, the power-law behaviour does not extend through the whole range of the degree distribution, rather for small values of k, the distribution is more similar to a random graph model. Garlaschelli and Loffredo [61], on the other hand, find a power-law only accounts for a small part of the whole degree distribution, and is cut-off at high values of k [62]. Furthermore, Fagiolo *et al.*[45] find power-law behaviour only in

www.manaraa.com

the middle of the distribution. Looking at individual commodity markets, Wu and Guclu [153] find the distributions of both export and import degrees exhibit inverse exponential decay, rather than power-law behaviour. Finally, in the bipartite case, the degree distribution for importers and exporters (when averaged over all commodities) follow heterogeneous patterns [52]. While the degree distribution, in the case of exports is approximately a power-law, the import degree distribution is approximated better by a truncated power-law in the form $\mathbb{P}(k) \sim k^{-\gamma} e^{-k/k}$ [52]. While Fernandez *et al.* [52] offer an interesting explanation as to why the two degree distributions differ (see Section 3.2.3 above), no explanation is offered as to why the network is formed by these two specific distributions.

3.3.2 Degree-Degree Correlations

Two other network statistic often discussed in the literature relate to degree-degree correlations across the network. The first is a measure of correlation between the in- and out-degree of each country in the network, and the second between a country's own degree, and the degree of its 'nearest-neighbours', meaning those countries that are in direct contact (see Chapter II). The first case only applies to aggregated trade networks, as in bipartite and disaggregated trade networks, importers tend not to export and vice versa [51]. Where applicable, Serrano *et al.* [122] found that the in- and out-degrees across the international trading system are highly correlated, with a coefficient of $r = 0.91$. This is perhaps not surprising given that, as noted above, the trade network is almost entirely symmetric, with 93% of edges being reciprocated [44]. This is to say, that for every export relationship (measured by the out-degree) between two countries, there is likely to be a reciprocal import relationship (measured by the in-degree) as well, and hence the correlation between the two.

The second case concerns the correlation between the average nearest-neighbour degree and degree $k_{nn}(k)$, which is effectively a measurement of the disposition of countries in the network towards trading with other countries in a similar position as themselves. Here, most authors find that vertex degree and average nearest-neighbour degree are negatively correlated, with caveats in some cases [43, 61, 62, 122, 128]. This is to say, that the international trading system is characterised by a *disassortative* pattern, whereby countries with many trade partners are typically connected with countries with few, and vice versa [128]. This is taken to imply a kind of hierarchy in the network, between the 'haves' (high-degree countries) and the 'have nots' (low-degree countries) [43]. Taken further, $k_{nn}(k)$ can be plotted to give a distribution showing the precise relationship between these two variables. Interestingly, Squartini *et al.* [128] find a

www.manaraa.com

highly linear trend, whereas, Fernandez *et al.* [52], and Serrano and Boguñá [122] both find a power-law decay in the form $k_{nn}(k) \sim k^{\gamma}$. In the latter case, $\gamma = 0.5 \pm 0.05$, but the power-law behaviour is only found in the tail of the distribution. Either way, this correlation appears to be highly stable over time, and does not change when measuring the network using either the directed or undirected approaches [43, 128]. The caveats to these findings come from analysis of weighted networks (which are discussed below) and when considering measuring the trade network of individual products. On the individual product level, the majority of markets still display power-law behaviour in $k_{nn}(k)$, however, in a few isolated cases, an *assortative* behaviour can is found, such that highly connected countries tend to associate with those that are alike [51]. A critique of this measure is offered by Squartini *et al.* [128], who compare the $k_{nn}(k)$ correlation between empirical trade networks and similar, but randomised, trade networks. They find that the $k_{nn}(k)$ correlations coincide, and argue that disassortativity is simply an effect of structural constraints of the network. They find the only way to remove assortativity, i.e., to have $k_{nn}(i)$ independent of $k(i)$, is to force it through a fixed generative mechanism. Naturally, these findings diminish the relative importance that should be placed on interpreting assortative/disassortative behaviours in international trade networks.

3.3.3 Clustering Coefficient

Along side the average nearest-neighbour degree, clustering coefficients are also thought to be an important indicator of the structure of international trade networks. The local clustering coefficient $Cl(i)$, represents the fraction of neighbours of a given country, which are also neighbours of each other [128]. The global clustering coefficient $\langle Cl \rangle$, is simply the average of the local coefficients across the network, and can be considered the probability that two neighbours of a randomly selected country are themselves connected to each other [13, 35]. On the global scale, trade networks are found to be far more clustered than would be expected, with $\langle Cl \rangle = 0.65$, approximately three times the values found in a random graph model of the same network [122]. The local clustering coefficients are found to be negatively correlated with country degree, meaning that trade partners of highly connected countries are poorly interconnected, whereas partners of poorly connected countries are highly interconnected [128, 43]. Serrano and Boguñá [122] finds that this relationship can be estimated by a power-law distribution in the form $Cl(k) \sim k^{-\gamma}$, with $\gamma = 0.7 \pm 0.05$. In the bipartite case, Fernandez *et al.* [52] find two different behaviours. For exporters, the bipartite clustering coefficient / degree correlation $Cl_{4b}(k)$, can best be fitted with a power-law distribution of the form $Cl_{4b}(k) \sim k^{\gamma}$, with the exponent $\gamma = -1.3 \pm 0.3$, across varying commodity markets. For importers, the value of the

www.manaraa.com

clustering coefficient decrease monotonously with the country degree [52].

Similar to the critique of average nearest-neighbour degree / degree correlations, Squartini *et al*. [128] compare $Cl(k)$ between empirical trade networks and randomised simulations to find that, again, the correlations are in very close agreement. This suggests that the clustering coefficient observed is simply explained by a constraint on the degree distribution, and finding a correlation between the two does not necessarily imply some higher order mechanism. The authors note that these results, taken together, highlight the importance of reproducing empirical degree distributions in network models of trade [128].

3.3.4 Strength Centrality

In weighted network representations of the international trading system, the strength centrality C_S of each country, and the distribution of these centralities, is an important property studied in the literature. Recall from Chapter II that the strength centrality is a measure of the sum weight of a vertex's edges, and can be interpreted as the total trade 'intensity' carried by each country [44]. Overall, the distribution of these values is found to be heavily right (or positive) skewed, such that the mass of the distribution is concentrated around low-intensity trading nations. Consequently, only a very few countries account for the majority of trade by volume, across the network [44]. Similarly, Bhattacharya et al. [21] found that the distribution of individual edge weights can be estimated by a log-normal distribution (i.e., also right skewed), such that we can say that a handful of edges account for the majority of trade across the network.

Researchers have also considered whether there is a correlation between a country's strength centrality and degree. Fagiolo *et al*. [44] find that $C_S(k)$ follows a general positive linear trend, such that highly-connected countries tend to be connected with stronger edge weights and vice versa. In the bipartite case, Fernandez *et al*. [52] find that this correlation follows a power-law pattern in the form $C_S(k) \sim k^{\gamma}$, where the exponent $1 \leq \gamma \leq 3$ depending on the commodity market sampled. Bhattacharya et al. [21] finds that there is also a correlation between a country's strength centrality and its GDP, in the form of a power-law, such that $C_S(i) \sim GDP_i^{\gamma}$, where $\gamma \approx 1.26$. However, this is not considered a particularity surprising result, as it simply confirms that countries with a higher GDP will be able to trade more than those with a lower GDP [21]. One final strength correlation studied is the average nearest-neighbour strength, which effectively measures the intensity of edge weights between the trade partners of a given country, and a country's own strength centrality. This relationship, given by $S_{nn}(C_S)$,

www.manaraa.com

is a measures of network strength assortativity/disassortativity [44]. In Fagiolo *et al*. [43], only a weakly disassortative relationship is found using this method.

3.4 Conclusion

Taking these results together, we can briefly summarise that the structure of the international trade network as follows. Countries in the network are highly heterogeneous across many properties. Trade in aggregate is highly reciprocal, but otherwise bipartite [51]. Connectivity is possibly distributed in a power-law fashion [19]. On the global scale, trade tends to be tightly clustered [122]. Trade across the network is characterised by two disassortative patterns. First, highly-connected countries typically trade with poorly-connected countries [128]. Second, the trade partners of highly-connected countries are, themselves, connected to poorly-connected countries [122]. However, trade networks are less disassortative when trade volumes are considers. Here, highly-connected countries tend to be connected by stronger trade links and, consequently, account for the majority of trade across the network [43].

www.manaraa.com

Chapter 4

Growth in Network Models of Commodity-trade Markets

4.1 Introduction

A ubiquitous feature in the trade network literature are reports of extremely high levels of heterogeneity in the degree distribution of the networks under study [19, 43, 54, 128, 129, 122, 21, 45]. These results, typically presented as power-law distributions, are taken to imply a so-called 'preferential attachment' model of network growth (see Chapter III for a detailed review). Recently however, serious critiques have been levelled at the methods used in these studies, and consequently, we may need to re-evaluate many of the conclusions drawn [32, 55, 130]. In this chapter, we outline the most common of these critiques and identify more appropriate methods for measuring the degree distributions of empirical networks. We also employ an alternative methodology that directly measures the growth in network connectivity. Using this method, we show that growth in many trade networks is constrained, resulting in a far different topology than the preferential model would imply. We develop a new single parameter model that reproduces many of the statistical features of empirical networks found across a wide array of commodity markets. The model can also reduce to a pure preferential growth model via tuning a single parameter, giving us the ability to see just how far reality falls short of this well-known theoretical ideal. This model could have wide reaching implications, both in terms of estimating future market growth and the robustness of critical markets. In the first part of this chapter, (Section 4.2), we present an in-depth commentary on the critiques of those methods commonly used in measuring the degree distributions of trade networks. In Section 4.3 we develop a new model of constrained growth and outline its main features and behaviours.

www.manaraa.com

Finally, in a concluding section (4.4), we discuss the implications of our findings for estimating future market growth and regarding the structural robustness of critical commodity markets.

4.2 Measuring Network Degree Distributions

4.2.1 Background

To put this discussion in context we first briefly reiterate and expand on the Barabási and Albert's [14] preferential attachment model (see Chapter II). The Barabási-Albert (B-A) model works by adding new vertices over time, that attach themselves *preferentially* to other highly connected (i.e., high degree, k) vertices already present in the network [15]. New vertices are thought to connect preferentially due to the advantages of being attached to other well-connected vertices. The model captures these features by the following two step process:

(a) Starting with a small number (m_0) of vertices, at every time step add a new vertex v_t with m edges (where $m \leq m_0$)—connected to the vertices already present in the system.

(b) When choosing the vertices to which the new vertex v_t connects, the probability Π that a new vertex will connected to an existing vertex v_i, depends on each vertices' connectivity k_i, such that

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}. \tag{4.1}$$

After t time steps, the model leads to a network with $N = t + m_0$ vertices and $m_0 + mt$ edges. Networks generated via this process are said to have a 'scale-free' topology, identifiable by a power-law degree distribution following,

$$\mathbb{P}(k) \sim k^{-\gamma}, \tag{4.2}$$

with the exponent γ (usually in the range $2 < \gamma < 3$), indicating the rate of decay. Note that γ is independent of the only parameter in the model, m the number of edges added at each time step [15].

While this model has support in the literature, evidenced by the numerous reports of power-law behaviour (see Chapter III), several authors have recently called into question many of the fundamental assumptions behind this body of work (not just trade networks in particular). Clauset *et al.* [32], Fox Keller [55], and Stumpf [130], have each sought to re-evaluate the

www.manaraa.com

evidence that has been presented and think critically about concluding that these networks are formed via the preferential attachment model, particularly when those conclusions are based solely on reports of power-laws. Their concerns stem from the fact that theories for power-law behaviour arose from large theoretical systems, as opposed to small real-world systems [130]. Specifically, in the case of trade networks, there are typically very few observations in the data, as the number of countries trading a specific commodity is limited (typically $50 < N < 150$). The small number of observations can cause large fluctuations in the tail of the network's degree distribution due to finite sample size effects [32]. This means that, even if a given network was formed by preferential attachment, the observed degree distribution would be extremely unlikely to exactly follow a power-law for all values of k, making it difficult to conclude whether a particular distribution is a power-law or not [130]. As a heuristic, we should expect that approximate power-law behaviour will be present in at least two orders of magnitude over both axes to be considered viable [130].

While these issues make verifying power-laws empirically a difficult procedure, numerous researchers compound the problem by neglecting to apply robust statistical procedures in their analysis [130], relying instead on only descriptive evidence from the network's degree distribution. We use Baskaran *et al.* [19] as an example and reproduce their results in figure 4.1, but it is common for power-laws to be confirmed in a similar fashion (see [43, 54, 128, 122, 21] for examples). Looking at figure 4.1, we see the frequency of degrees in various commodity markets as plotted histograms on doubly logarithmic (*log-log*) axes. Baskaran *et al.* [19] reason that by taking the logarithm of both sides of equation (4.2), we obtain

$$\log \mathbb{P}(k) \sim -\gamma \log k. \tag{4.3}$$

Accordingly, the log of $\mathbb{P}(k)$ is expected to depend linearly on the log of k. On a *log-log* plot, a power-law should therefore present itself as a straight line, with the slope given by $-\gamma$. While this method is certainly straightforward, Clauset *et al.* [32] points out that our human ability to discern patterns in data, even where there are none, can lead researchers to confirm power-laws in almost any case using this qualitative criteria. Indeed, Baskaran *et al.* [19] conclude that the plots in figure 4.1 provide sufficient evidence that the selected commodity markets should be considered as having a scale-free topology. Several researchers develop this method further, typically by adding a line of best fit and estimating the slope (and thus the exponent $-\gamma$), by performing a least-squares linear regression (see [51, 52] for examples). This also gives standard errors for the slope, and the fraction r^2 of variance accounted for by the fitted line,

www.manaraa.com

which is taken as an indicator of the quality of the fit [32]. However, Clauset *et al*. [32] note that this method[1] is highly problematic and can suggest power-law fits to data that are known to be non power-law distributions.

4.2.2 Methods

A more robust statistical method for testing whether or not a power-law hypothesis is plausible, is described at length in Clauset *et al*. [32]. They recommend using the Kolmogorov-Smirnov (KS) statistic [50], applied to the empirical distribution function (herein CDF) of the data as a *goodness-of-fit* measure and to generate p-values. The KS statistic is typically used to quantify the distance (sensitive to both location and shape) between two probability distributions. In this case, the supremum (greatest) distance D_n can be calculated by

$$D_n = \sup_x |F_n(x) - F(x)|, \tag{4.4}$$

were $F_n(x)$ is the EDF of the n observed vertex degrees, and $F(x)$ is the cumulative distribution function (CDF) of the hypothetical power-law model that best fits the data. The hypothesis for this test would be as follows:

- H_0: The data follow a power-law distribution.
- H_1: The data do not follow a power-law distribution.

If the data is a perfect fit, the KS statistic, $D_n = 0$. Larger values of D_n indicate a poor fit. As such, the null hypothesis is to be rejected if D_n is greater than some critical value (found by a lookup in a KS table) determined by the level of statistical significance required. Here, we consider the result to be statistically significant where $p > 0.05$ in the Kolmorogov-Smirnov test. Note that the KS tests assumes independence between the distributions under consideration.

As noted above, finite sample size effects can cause large fluctuations in the tail of the degree distributions [32]. We should, therefore, only expect a power-law to fit above some lower bound, $k_{\min}$. Determining this point and discarding all data for $k < k_{\min}$ becomes very important to obtain accurate results using the KS test. There are several ways to approximate the correct value of $k_{\min}$, including using Bayesian information criterion (BIC) for example. However, in an earlier paper, Clauset *et al*. [33] develop a method that re-uses the KS statistic outlined above, thereby, reducing the complexity of the analysis. This method involves simply

[1] Technical details of these critiques are described in the appendix of Clauset *et al*. [32].

www.manaraa.com

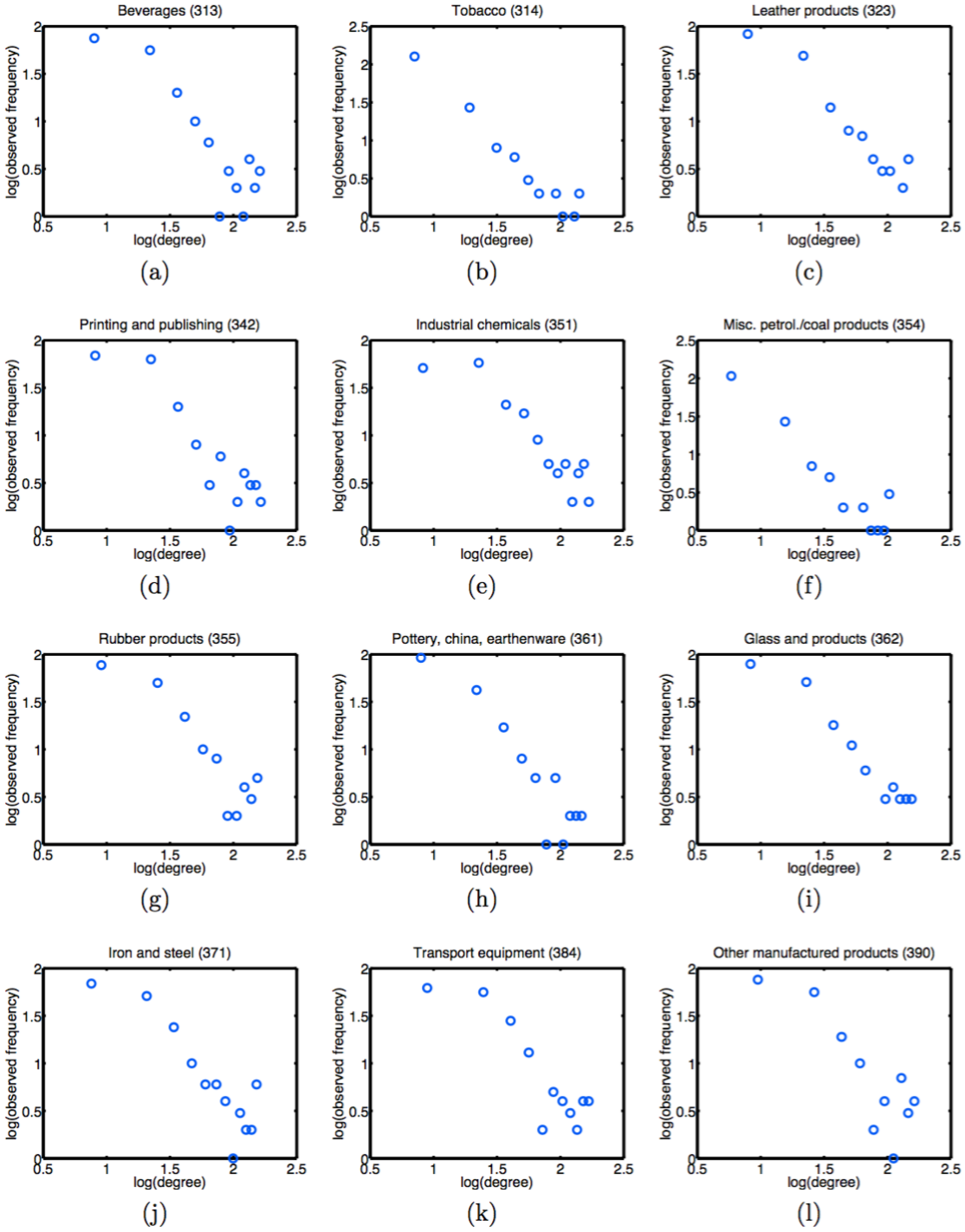


Fig. 4.1 Empirical degree distributions for several commodity markets in 1980. Note the *log-log* axis. Reproduced from Baskaran *et al.* [19, p.14, fig.1].

www.manaraa.com

applying the KS statistic to a model fitted to data in the region $k \geq k_{\min}$, for each value of $k_{\min}$ and selecting that value of $k_{\min}$ that minimises D_n. To understand this concept visually, one might imagine trying to find the point beyond which, a somewhat concave, degree distribution becomes a straight line on a *log-log* plot.

Even if a distribution is well fitted by a power-law using the method above, it does not necessarily prove a power-law is the correct distribution for the data. There may be several other heavy-tailed distributions that match the data equally well, if not better [32]. Failing to consider these alternative distributions is a large weakness in the current trade network literature, which is perhaps due to the overwhelming prominence of the Barabási-Albert model [14], which is undoubtedly the most cited work in the field. Outside of trade network research, it is more well known that many different models can result in the same network structure being produced [55]. Thus, finding a particular distribution does not necessarily imply a specific mechanism giving rise to it [130]. In fact, researchers have shown that a scale-free network topology can be generated by a number of different models, with and without the need for a preferential attachment mechanism [88, 58, 154]. In addition, Barabási himself outlined several other network growth models that do not result in power-law degree distributions [14]. For example, in a lesser known variant of the A-B model, preferential attachment is replaced with a *uniform* attachment mechanism. In contrast with the A-B model, the resulting degree distribution follows an exponential form, which shows as a concave downward curve on a *log-log* plot (or more appropriately as a straight line on a *lin-log* plot). Barabási [14] outlines the model as follows:

(a) Similar to the A-B model, start with a small number (m_0) of vertices, and at every time step add a new vertex v_t with $m(\leq m_0)$ edges

(b) The new vertex v_t connects with equal probability to the v_i vertices already present in the system, independent of their degree k_i, such that

$$\Pi(k_i) = \frac{1}{(m_0 + t - 1)}. \tag{4.5}$$

As noted the degree distribution follows the exponential form,

$$\mathbb{P}(k) = B \exp(-\beta k), \tag{4.6}$$

www.manaraa.com

with the coefficients[1] given by

$$B = \frac{e}{m}, \qquad \beta = \frac{1}{m}. \tag{4.7}$$

While their is no economic theory to imply that this is a more plausible model of trade network growth, it is useful to be aware of other possibilities, when judging the results of any statistical analysis.

4.2.3 Results

If one follows the method outlined above, selecting the best model among a given set of plausible models is relatively straightforward. We need only apply the same KS statistical test to each alternative model, fitting the appropriate distributions, and use the resulting p-values to rule out models that do not meet the level of statistical significance. If we are left with several statistically significant alternatives, we can use the goodness-of-fit measures to make a judgement as to which has the strongest case. So for example, if among a set of statistically significant distributions, the distance D between a power-law fit and the data is low, while the distance for the remaining distributions is high, then the case for a power-law fit is strengthened.

We apply this model selection method to data from the UN Comtrade dataset [141], measuring the degree distribution of the top ten internationally traded commodities[2]. From the database, we select annual transaction data over the period 1992–2014, which is the longest uninterrupted period of data available. Trade data prior to 1992 does exist, however, the dissolution of the Soviet Union (officially on December 25, 1991) makes comparing the network before and after the event rather contentious. The Soviet Union split into the Russian Federation and fourteen other post-soviet states[3], leading to a large jump in the number of trade relationships with those who previously imported from—or exported to—the Soviet Union. This event also ended the era of large trading blocks, which had effectively divided the developed world in two for several decades prior. While understanding the effect of this event on the network structure of trade would be an interesting topic on its own, for our purposes, combining trade data from the period before and after, does not seem like a sensible idea as this event might not be considered part of the natural evolution of the network. The data selected are used to construct a unweighted-directed adjacency matrix of a network, where vertices represent

[1] where e refers to 'Euler's number', i.e., the base of the natural logarithm.

[2] The top ten internationally traded commodities used in this study; Petroleum Oil, Natural Gas (Liquefied), Coal, Wheat and Meslin, Maize (Corn), Coffee, Cotton, Cane and Beet Sugar, Silver, and Gold.

[3] Post-soviet states include: Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan

www.manaraa.com

countries, and edges denote the import/export of a commodity between those countries. The Matlab function to do this is reproduced in Appendix A1.

The results of our statistical analysis of the Comtrade data are shown in tables 4.1 and 4.2. The most notable result is that none of the degree distributions, for either imports or exports, can be considered a plausible power-law distribution (see the first set of columns in table 4.1 and table 4.2). An inspection of the model fits in figure 4.2 or figures A.1 and A.2 in Appendix A3 show that a power-law fit would probably not even be considered, if one were to approach these data without the Barabási-Albert (B-A) model in mind. Four of the export markets and five of the import markets are plausible exponential distributions (see the second set of columns in table 4.1 and table 4.2), with Cotton and Crude oil being very strong fits. The third model fit tested is an exponentially truncated power-law in the form

$$\mathbb{P}(k) \sim k^{-\gamma} e^{-\beta k}. \tag{4.8}$$

This model is essentially a combination of the latter two models, which typically presents itself on a *log-log* plot as a straight line that tapers into a downward concave curve in the tail of the distribution. We find this model is a plausible fit to nine out of ten of the selected export markets but only three import markets (see the third set of columns in table 4.1 and table 4.2).

These results shed some light on a particular argument made in the literature. Specifically, that importers and exporters differ in their decision making process, and consequently play different roles in the network's topology (see Fernandez *et al*. [51, 52]). If this theory is true, we should find a marked difference in the resulting degree distributions of import and export markets. While it is fair to say that exports markets are likely to be generated by a different process (i.e., one that likely leads to a truncated power-law) to the import markets, we find no degree distribution is strictly unique to importers or exporters. Therefore, our results run counter to the assertions in [52]. Note that several import markets[1] do not fit any of the three models selected. Having tested several other distributions (Weibull and Log-normal for example), we found none that result in better model fits than those given by the exponential model. Given the hundreds of other commodity markets we could test, and the number of possible models, it is doubtful we would find a single distribution to fit all markets.

[1] Specifically; Coal, Maize, Natural gas, Sugar, and Wheat

www.manaraa.com

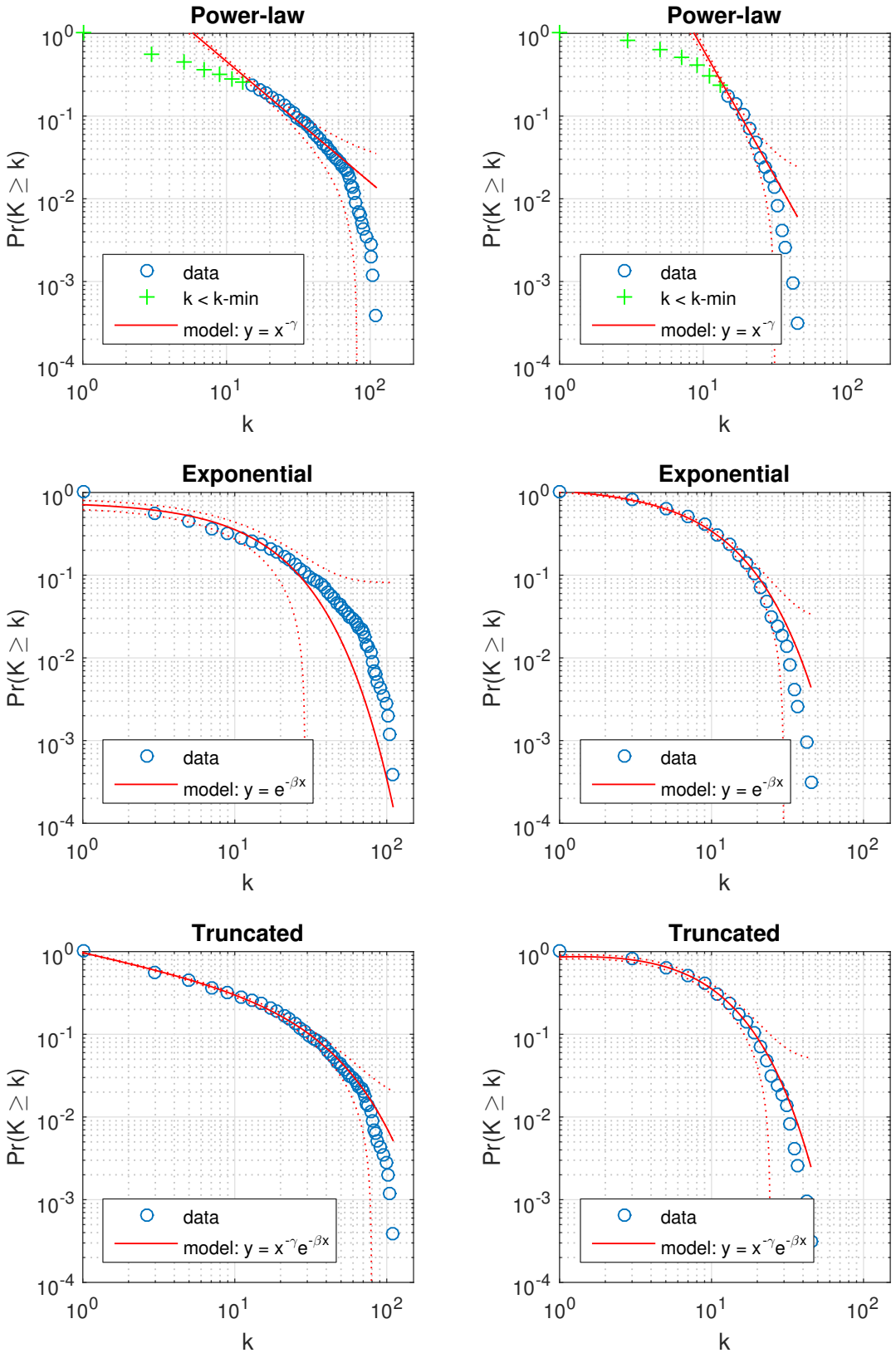


Fig. 4.2 Empirical degree distributions (exports down the left, imports down the right) for the wheat market. Data source: [141].

www.manaraa.com

	Power-law				Exponential			Truncated			
Market	p	D	$k_{\min}$	γ	p	D	β	p	D	γ	β
Coal	0.00	0.38	34	-3.02	0.02	0.23	-0.09	**0.24***	0.15	-0.40	-0.03
Coffee	0.00	0.35	50	-2.73	**0.84***	0.07	-0.04	0.02	0.18	-0.19	-0.03
Cotton	0.00	0.31	24	-3.03	**1.00***	0.05	-0.08	**0.18***	0.18	-0.20	-0.05
Oil	0.01	0.34	19	-2.90	**1.00***	0.06	-0.12	**0.55***	0.15	-0.18	-0.08
Gold	0.00	0.38	3	-1.02	0.00	0.39	-0.15	**0.48***	0.14	-0.48	-0.05
Maize	0.00	0.34	37	-2.64	0.00	0.31	-0.08	**0.97***	0.07	-0.34	-0.03
Gas	0.01	0.25	27	-1.78	0.00	0.32	-0.08	**0.48***	0.11	-0.31	-0.04
Silver	0.02	0.25	22	-1.76	**0.17***	0.17	-0.07	**0.56***	0.12	-0.35	-0.03
Sugar	0.02	0.20	27	-1.52	0.00	0.30	-0.06	**0.40***	0.11	-0.28	-0.03
Wheat	0.00	0.38	39	-2.52	0.01	0.23	-0.08	**0.33***	0.13	-0.39	-0.03

Table 4.1 Degree distribution model fits for selected export markets. Statistical significance is indicated by (*), where $p > 0.05$ in the Kolmorogov-Smirnov test. Data source: [141].

	Power-law				Exponential			Truncated			
Market	p	D	$k_{\min}$	γ	p	D	β	p	D	γ	β
Coal	0.00	0.50	12	-2.80	0.00	0.38	-0.13	0.00	0.38	0.07	-0.13
Coffee	0.00	0.40	43	-2.81	**0.21***	0.13	-0.06	0.00	0.68	-52.34	2.46
Cotton	0.00	0.39	2	-0.75	**0.62***	0.12	-0.04	**0.48***	0.14	-0.26	-0.04
Oil	0.00	0.38	4	-1.09	**0.61***	0.14	-0.07	**0.43***	0.16	-0.23	-0.07
Gold	0.04	0.24	20	-2.03	**0.21***	0.17	-0.09	0.00	0.97	-52.99	2.49
Maize	0.00	0.42	15	-2.89	0.00	0.33	-0.14	0.00	0.43	0.27	-0.14
Gas	0.00	0.42	16	-2.96	0.00	0.30	-0.13	0.00	0.40	0.29	-0.13
Silver	0.00	0.37	19	-2.83	**0.05***	0.23	-0.08	**0.15***	0.19	-0.05	-0.08
Sugar	0.00	0.39	23	-2.80	0.01	0.24	-0.09	0.00	0.35	0.24	-0.09
Wheat	0.00	0.49	14	-3.08	0.00	0.39	-0.15	0.00	0.45	0.20	-0.15

Table 4.2 Degree distribution model fits for selected import markets. Statistical significance is indicated by (*), where $p > 0.05$ in the Kolmorogov-Smirnov test. Data source: [141].

www.manaraa.com

4.3 A New Model of Constrained Network Growth

4.3.1 Methods

Considering the failure of these methods to find a comprehensive model and fit that could explain all markets selected, we apply an alternative method that directly measures the growth in network connectivity over time, rather than the resultant degree distribution which is only an artefact of that process [114]. This method was introduced by Barabási himself in one of his earlier papers [15], but seems to have gone unused in the network literature, bar one notable exception [114]. Connectivity growth is controlled by the edge-attachment rate A_k, which measures the rate at which edges attach to a vertex of degree k. Attachment rates in networks generated via preferential attachment are found to grow asymptotically linear in k (examples include citation networks and the world wide web [114, 15]). That is to say, the rate at which new edges attach to vertices, will correlate positively with the vertex degrees. To measure the empirical attachment rate in commodity trade networks, we first measure the degree of each country (vertex) in the first year available, then count the number new bilateral trades (edges) relating to that country over the following year. We repeat this process for each year 1992–2014, and for each of the selected markets. By aggregating these data, we can calculate the overall yearly attachment rate.

The results of this method are shown in figure 4.3. We can see that the attachment rate A_k increases approximately linearly in k, where $k < 50$, but levels off in the tail of the distribution. Consequently, the attachment rate significantly under-performs our expectations based on a preferential attachment mechanism. For example, we would have expected empirical observations of A_k to be 2-3 times higher in the tail of the distribution (e.g., $A_k \sim 60$, where $k = 125$), if the Barabási-Albert model [14] were being used to generate these networks.

In addition, we find that edge attachments can form both *intrinsically*, via edges linking existing market participants, and *extrinsically*, from edges connecting new entrants to the market. This distinction appears to have gone unnoticed in the literature, perhaps on account of Barabási's formulation [15] only allowing for the latter. This is not necessarily a criticism, as extrinsic network growth models can be very apt in many cases. Guillaume and Latap [72] for example, study a Film/Actor network, which connects actors together who star in the same movies. Accordingly, the attachment of new edges (between actors) in this network is always contingent on the creation of a new vertex (the release of a new film). However, this growth model is not applicable to commodity-trade networks, or possibly for trade networks in

www.manaraa.com

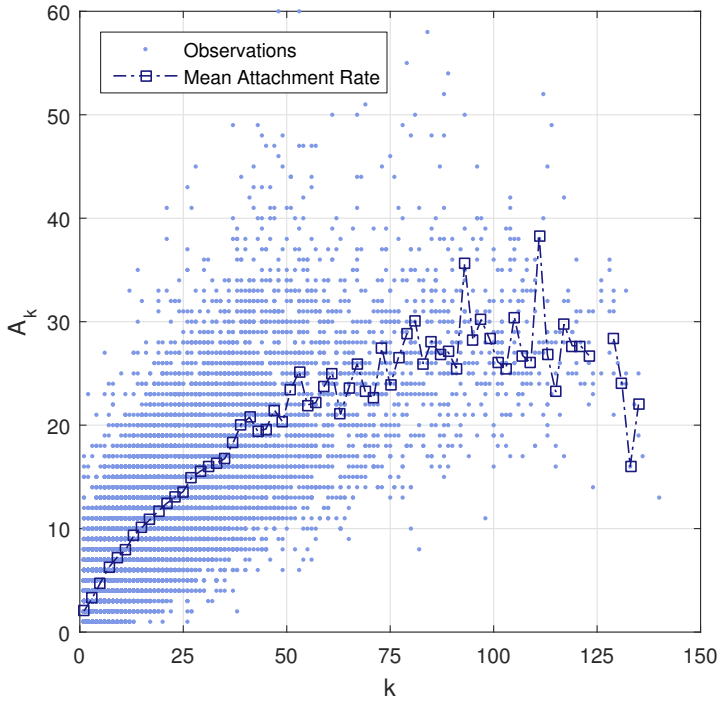


Fig. 4.3 Empirical edge-attachment rate A_k relative to vertex degree k. Data source: UN Contrade [141].

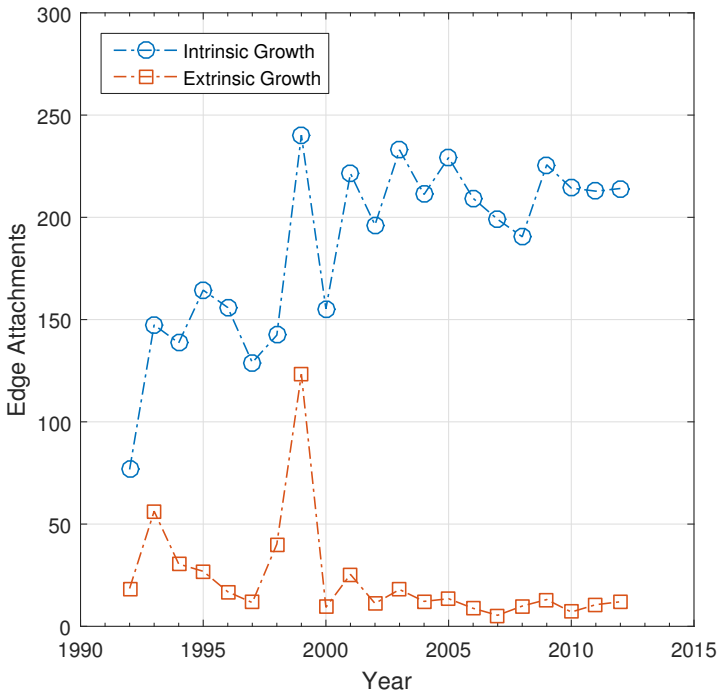


Fig. 4.4 Edge-attachments by origin: an intrinsic edge is attached between two pre-existing vertices, and an extrinsic edge is attached to a new vertex. Data source: UN Contrade [141].

www.manaraa.com

general. Here, two pre-existing vertices can attach together *independently*, and do so often. In fact, looking at the origins of edge-attachment events in figure 4.4, we find that intrinsic edge attachments greatly outnumber extrinsic attachments, and that this disparity appears to be only growing over time.

4.3.2 The Model

We incorporate the difference between intrinsic and extrinsic growth, in a new generative model of *constrained* network formation. Central to this model is the parameter λ, which measures the probability on a scale $0 \leq \lambda \leq 1$ of a new edge being generated extrinsically via the addition of a new vertex, as opposed to connecting pre-existing vertices intrinsically. Changing this one parameter can constrain extrinsic growth in the model in-line with empirical based estimates.

Briefly, this new model is outlined as follows. A graph is formed by $G = (V, E)$, where V is a set of vertices, and E a set of edges between them. Time t refers to number of $t-$edge attachment events. At time t_0, a small seed network is formed of two vertices, connected together by an edge. The model then proceeds to evolve the graph each time step, as follows:

(a) *Extrinsic Growth:* With probability λ, add a new vertex v_t and attach it to a pre-existing vertex $v_i \in G$, chosen via preferential attachment, where the probability of attaching e_t to vertex v_i is relative to its connectivity k_i, such that

$$\Pi(k_i) = \frac{k_i}{\sum_{j \in G} k_j}. \tag{4.9}$$

(b) *Intrinsic Growth:* Otherwise, attach an edge e_t between two pre-existing vertices in G, starting at a random source vertex and attached to a target vertex chosen via the preferential attachment mechanism in equation (4.2).

Note that in step (b), edges can only be connected once. If the model does select a duplicate edge, another target vertex is chosen following equation (4.2). If all possible edges are connected, the model returns to step (a).

The bipartite case

Commodity markets are bipartite in nature (see Chapter III), with the distinction being between sets of importers and exporters. We account for this fact by developing a bipartite version of the model, which we outline below. However, it should be noted that the only distinction

www.manaraa.com

between each set in this version of the model is the value of λ, which can take different values for importers and exporters. As such, we can model either set independently using the simpler unipartite model, and do so for modelling the results later in this section.

The bipartite model is outlined as follows. A bipartite graph is formed from the triple $G = (U,V,E)$, where U and V are two disjoint sets of vertices, and $E \subseteq U \times V$ is the set of edges between them. Time t refers to number of $t-$edge attachment events. At time t_0, a small seed network is formed of at least two vertices, one in each set, connected together by a single edge. The model then proceeds to evolve the network over time via two independent mechanisms:

(a) *Intrinsic Growth:* At every time step t, attach an edge e_t between two pre-existing vertices in G, chosen via preferential attachment [15]. Specifically, the probability of attaching e_t to vertex $v_j \in V$ is relative its connectivity k_j, such that

$$\Pi(k_j) = \frac{k_j}{\sum_{i \in V} k_i}, \tag{4.10}$$

and similarly, the probability of attachment to vertices $v_i \in U$ is relative to k_i.

(b) *Extrinsic Growth:* With probability λ we add a new vertex $v_i \in U$, attached to a pre-existing vertex $v_j \in V$ via preferential attachment outlined in equation (4.10). Similarly, a vertex is added in V with probability λ and attached to U via preferential attachment. The parameter λ measures the approximate rate at which vertices are added to the graph relative to the number of $t-$edge attachment events. So for example, where $\lambda = 1$, one new vertex is added for every edge. Where $\lambda < 1$, extrinsic growth is constrained.

Edges in the bipartite model can only be connected once. The selection of duplicate edges are dealt with by a similar process to that outlined above in section 4.3.2.

4.3.3 Main Results

The main feature of the model is that increasing the constraint on extrinsic growth has the effect of homogenising high-k vertex degrees. With increasing degree homogenisation, the probability of a new edge attaching to higher degree countries is reduced, effectively limiting how 'preferential' the attachment mechanism can be. This results in attachment rates becoming increasingly constrained as λ tends to zero. An interesting dynamic emerges between the

www.manaraa.com

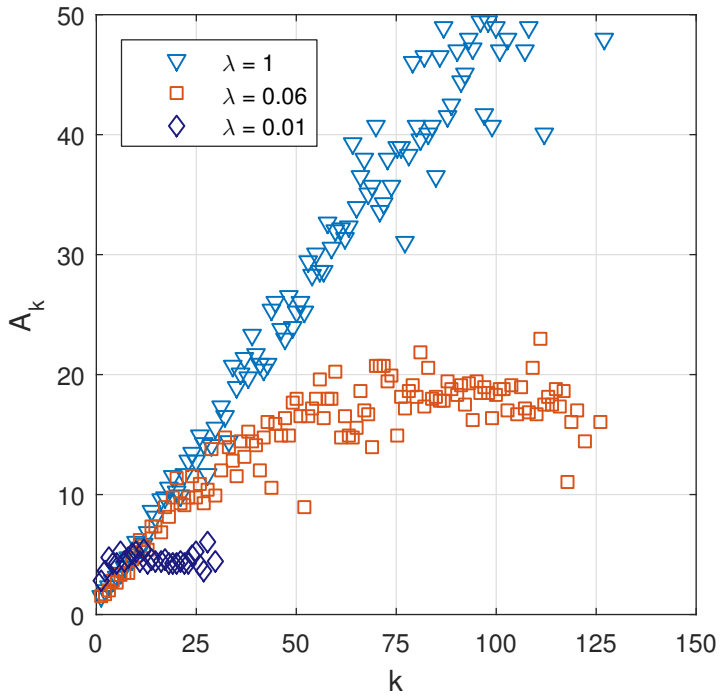


Fig. 4.5 Model edge-attachment rate A_k relative to vertex degree k.

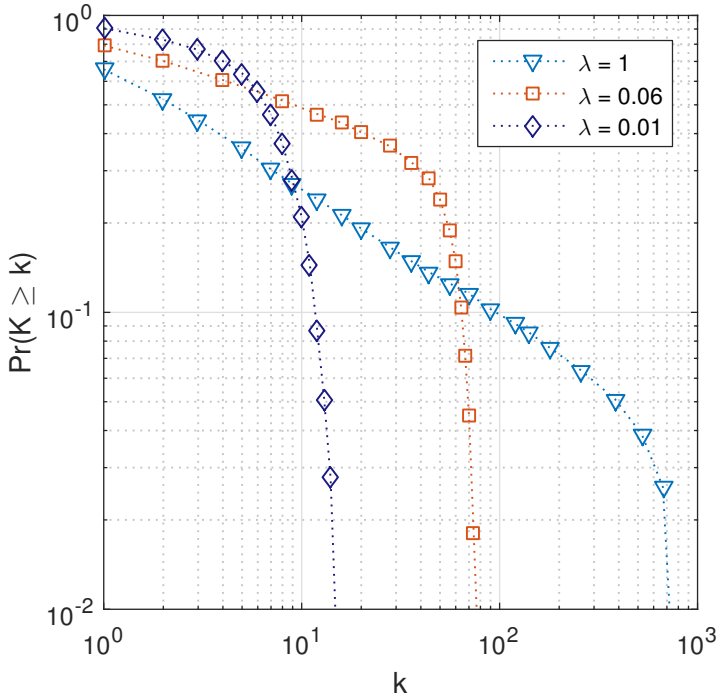


Fig. 4.6 Model degree distribution, showing cumulative probability relative to vertex degree k (note *log-log* axis).

www.manaraa.com

parameter limits of the model ($0.01 < \lambda < 1$), where the resulting network displays a core/periphery like structure, with the core connected in a uniform fashion, and the periphery connected preferentially. Alternatively, at the parameter limits of the model, where $\lambda = 1$, the resulting networks are similar to those produced using the Barabási-Albert (B-A) model [15]. Here, all edges attach extrinsically and the network develops into a scale-free topology, with A_k linear in k and a power-law like degree distribution. Setting $\lambda = 1$, effectively executes the model code as one would programme an B-A model, so it is no surprise that they would produce the same outcome. At the other extreme, setting $\lambda \leq 0.01$ suppresses the preferential attachment mechanism to such a degree that the model produces networks functionally equivalent to a random network. In simulations, the attachment rates are uniform and the resulting degree distribution is exponential. This leads to networks similar to the alternate uniform A-B model outlined above, or to the Erdös-Rényi (E-R) [41] random network model outlined in Chapter II. The results of our simulations can be seen in figure 4.5 and figure 4.6. The results shown are calculated by averaging over 1000 realisations of the model with choice of parameter λ. In addition, a breakdown of the model parameters settings and observed behaviours from the simulations are shown in Table 6.1.

Table 4.3 Observed Model Dynamics in Simulations.

Parameter	Edge Attachment	Network Structure	Degree Distribution
$\lambda = 0$	No Growth	–	–
$\lambda \leq 0.01$	Uniform	Random	$\mathbb{P}(k) \sim e^{-\beta k}$
$0.01 < \lambda < 1$	Constrained	Core/Periphery	$\mathbb{P}(k) \sim k^{-\gamma} e^{-\beta k}$
$\lambda = 1$	Preferential	Scale-free	$\mathbb{P}(k) \sim k^{-\gamma}$

Empirical Estimates of Model Parameters

For the model to be of use in predicting future market growth, we first need to derive empirical estimates of the model parameters. To do this, we find the ratio of intrinsic-to-extrinsic edge-attachment events in each market, for each of the 22 years studied. The distribution of these ratios gives an estimate of the parameter λ across a wide range observations. The results are shown in figure 4.7, where we can see a distribution of observed parameter values. The most notable feature of this distribution is that in most years $\lambda \sim 0$, meaning that in those

www.manaraa.com

year, edges are only attaching intrinsically between existing vertices. The distribution then decays very rapidly, with the highest parameter value $\lambda = 0.06$. It becomes clear that rate of extrinsic growth in real world networks fall far short of the ideal set out in the A-B model, where we would expect the parameter value to be closer to $\lambda = 1$. The limiting factor on edge attachment rates in real-world trade networks is likely due to protectionist tariffs on imports and large subsides given to leading exporters, both of which create barriers to entry for new market participants. There is another possibility however, in that the trade networks could have grown much faster in past, but since plateaued. Unfortunately, the data available does not cover the full growth cycle of each commodity market, rather a period of time after maturity.

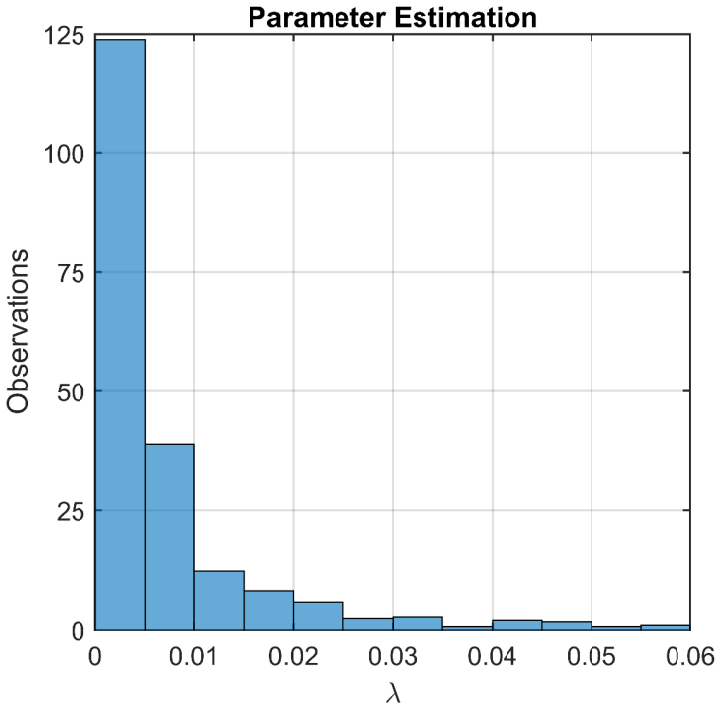


Fig. 4.7 Estimation of model parameter λ from empirical data. Data source: [141].

Model validation

To test whether the constrained growth model can reproduce the growth dynamics of real-world commodity-trade networks, we compare edge-attachment rates from simulations with empirically estimated parameter values taken from Figure 4.7, against the historical attachment rates shown in Figure 4.3 (*a*). In these simulations, the model is first run for a relatively long period of time (t_1), until the network structure has had a chance to mature. We then record

www.manaraa.com

the degree $k(t_1)$ of each vertex, before running the model again for a further period of time (t_2), but one much shorter than before. After this, we record the number of new edges attached to those vertices of k degree between periods t_1 and t_2, i.e., $A_k = k(t_2) - k(t_1)$. Increasing t_1 lets the network grow further—increasing the range of degrees, whereas, the length of time between t_1 and t_2 changes the maximum edge attachment rate—but importantly, not the shape of the distribution. In the empirical data, each period of time is measured in years, whereas, in the model, time is measured but t-edge attachment events. Therefore, the time between t_1 and t_2 needs to be set to match approximately one year of empirical data. We use a two-sample Kolmogorov-Smirnov (KS) test to see whether the distributions generated by our constrained growth model can reproduce those recorded in real-world commodity-trade networks[1]. Similar to the one-sample test described in Section 4.2.2 the greatest distance $D_{n,n'}$ between the model sample n, and the historical sample n' can be calculated by

$$D_{n,n'} = \sup_x |F_{1,n}(x) - F_{2,n'}(x)|, \tag{4.11}$$

where $F_{1,n}(x)$ is the EDF of the model attachment rate, and $F_{2,n'}(x)$ is the EDF of the historical attachment rate. Using this test we find a distance between the two samples $D_{n,n'} = 0.19$, which is found to have a moderately statistically significant p-value of 0.07 (where significance is defined as $p > 0.05$). The comparison between attachment rates was made by setting $\lambda = 0.06$, $t_1 = 2000$, and $t_2 = 2500$. These results show how the model is able to reproduce empirical edge attachment rates.

4.4 Conclusion

The main application of our findings is in predicting the future growth of commodity markets, using the constrained growth model to estimate both size and connectivity. For example, a Monte Carlo simulation method could be employed to give bounded projections of the future market structure at set time intervals as required. In the Monte Carlo simulation, the current market structure of a given commodity market would be used as a seed network from which the constrained growth model would be run. The model parameter values for each specific market can be estimated based on past data in a similar manner to the methods used in Section 4.3.3. Each run of the simulation would then grow the network for a specified period of time. After running the simulation many times, the statistical properties (degree rankings, degree distribu-

[1] The edge attachment rate from the empirical data is averaged over all commodity markets studies, whereas the model edge attachment rate is averaged over 100 realisations

www.manaraa.com

tion, total number of vertices etc...) can be averaged across the resulting networks to give a predicted value.

Using this procedure with the Barabási-Albert (B-A) model results in a 1600% increase in network size (N vertices) and a 1000% increase in edge density (as defined in Chapter II) compared to the constrained growth model with empirically estimated parameters. Therefore, using the B-A model can predict highly differing, and possibly unrealistic results. Future work could focus on further validating and improving the constrained growth model's performance in this area.

Another major implication of our results regards the structural robustness of critical markets, particularly the agriculture and energy sectors. Robustness here, is defined as a network's ability to withstand perturbation[1]. In a network setting this typically involves the removal of specific vertices from the network [5]. Vertex removal makes it increasingly difficult for the remaining vertices to interact with one another, and will eventually leave some vertices unreachable [4]. Real-world examples of these perturbations include, extreme weather events, trade restrictions, local conflicts, and piracy. Each of these leads to market fragmentation and higher prices. In the literature, there are two general categories of perturbations [153];

- *Error*: Vertices are removed at random, representing environmental factors.
- *Attack*: Specific vertices are removed, based on their connectivity.

Researchers have compared the robustness of the two network models discussed in Chapter II, namely, the Barabási-Albert (B-A) preferential attachment model [14], and the Erdös-Rényi (E-R) random graph model [41]. They find that the scale-free networks produced by the B-A model display an unexpected level of robustness to errors, but are extremely vulnerable to intentional attacks. On the other hand, random networks are not as robust to error, but are more resilient to intentional attacks [5, 72, 34]. The response of these models to vertex removal varies due to the underlying network degree distributions [4]. Fortunately, most real-world examples of perturbations fall in to the *error* category. While some events, such as embargoes and blockades are 'targeted' in a sense, as they are aimed at specific countries, they are not targeting the network based on connectivity (this tactic might be a possibility in the future, however), and so are better modelled as random errors.

[1] Defined as a deviation of a system, moving object, or process from its regular or normal state or path, caused by an outside influence [99].

www.manaraa.com

As noted above, use of the constrained growth model with empirically estimated parameters produces network structures that differ significantly from the scale-free networks generated using the Barabási-Albert (B-A) preferential attachment model. In particular, network degree distributions are far more homogeneous when using the constrained growth model, which makes them less likely to be robust to error type perturbations. Future work could be done to quantify the difference in robustness between the constrained growth model and it's alternatives.

www.manaraa.com

Chapter 5

Price Variation in Network Models of Commodity-Trade.

5.1 Introduction

International trade is a major source of income for many developing countries and plays an essential role in global economic development [139]. However, restrictions on trade introduce asymmetries in the economic interactions between consumers in global markets [126]. These asymmetries can lead to varying prices for the same goods simply due to the network structure [86]. While having to pay higher prices is detrimental for low-income consumers, low prices can prove challenging for producers as well. In this chapter, we apply a network model of commodity markets to measure the effect of the network structure on price asymmetries. A simple economic model is used, whereby standard price formation factors are specified so that any variation in commodity prices can be ascribed solely to variation in the underlying network structure. Using this model we find that exporters are able to secure favourable prices simply due to their position in the network. We also look at how global prices vary between networks with different topological properties. Our findings combine theoretical analysis, modelling, and applications to real-world trade data, the results of which are interpreted in terms of both economic theory and international trade policy.

The chapter is structured as follows: In section 5.2, we give some background on the standard economic model from which much of the literature is based. In section 5.3, we introduce the graphical exchange model developed by Kakade *et al.* [86]. Following the model outline, we investigate the determinants of local price variation in network-like markets in section 5.4. In the results section 6.6, we present the output of the model, first in terms of

www.manaraa.com

determining the theoretical upper and lower bounds of price variation (section 5.5.1), and second, we apply the model to empirical datasets representing real-world commodity markets (section 5.5.2). Section 6.8 concludes with an overview of the study.

5.1.1 Background

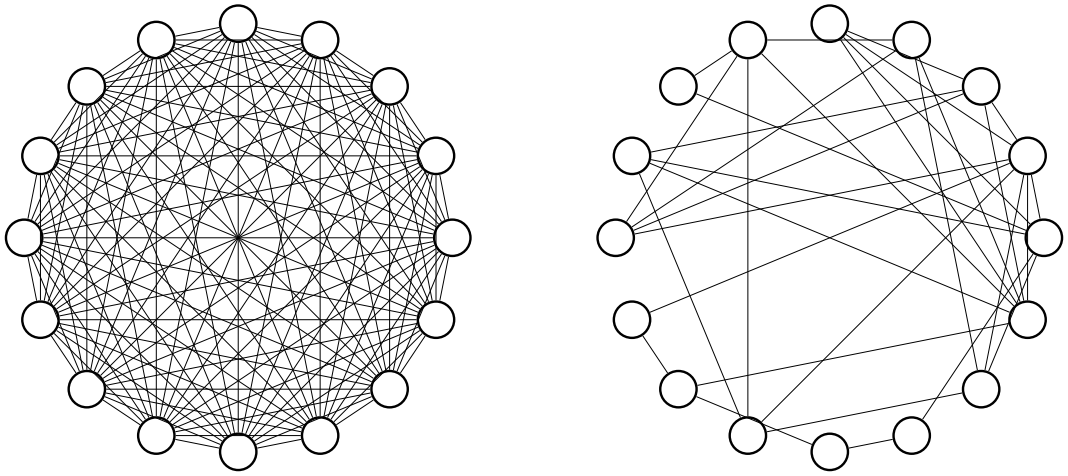
As noted in the introduction chapter to this thesis (Chapter I), economists have a long history of developing mathematical models of markets and trade. However, generally speaking, these models do not explicitly consider the network-like structure of markets. This is largely due to the simplifying assumption of 'perfect' competition, that was introduced in the nascent years of economic modelling (the Fisher [53] model of competitive markets was introduced in 1926, for example). Perfect competition is a theoretical market structure, defined by several conditions that have proven highly controversial in the development of general equilibrium theory [9]. The relevant assumption to this discussion, is that under perfect competition, no restrictions can be placed on trade[1]. This condition excludes the possibility of consumers extracting economic rents (i.e., unearned revenue) due to monopolistic, or monopoly-esque situations. As such, in a perfectly competitive market, prices are formed based on the aggregate balance of supply and demand across that whole market. In this case, it is often said that prices are formed 'globally', and that there exists a global price for the good. Real-world markets differ from this model in many ways, but most relevant to the network perspective, is that markets are typically characterised by restricted trade, rather than perfect competition [9]. Restricted trade can be due to the lack of a negotiated free-trade agreement between nations, or a unilateral government policy placing restrictions on imports/exports. Alternatively, these restrictions could be due to physical causes, such as natural disasters, or limited shipping routes. Restrictions on trade introduce asymmetries into the market's structure, meaning that some consumers have limited options with who to trade, and giving others a monopolistic-esque advantage—through which they can extract higher prices. This structural asymmetry ultimately leads to variations in price (or prices) between consumers and across the market as a whole. The concept of a 'global' price or global price formation, is thus replaced by local prices, determined, in part, by each consumer's position in the market structure.

The distinction between a 'perfect' and restricted market can be made clearer by looking at the situation from a graphical perspective. By modelling a market as a network, representing each consumer as a 'vertex' or node, we can draw an edge between two consumers where trade

[1] The other conditions being: large n consumers, perfect information, homogeneous goods, and no price makers [9].

www.manaraa.com

is permitted. Then, where no edge exists, we know where trade is restricted. Like most trade network models (see Chapter III for examples), these networks are typically used to represent a global market, where each vertex represents a nation, and each edge represents bilateral trade between nation pairs. But the network could just as easily represent a social network of individual businesses, with edges representing existing business relationships between clients [85]. In figure 5.1, we draw two graphical representations of a simple market model where consumers can act as both buyers and sellers. The first panel (*a*), represents a perfect competitive market under free-trade, and the second panel (*b*), represents a market subject to restrictions on trade.



(a) Perfect Market (b) Restricted Market

Fig. 5.1 (a) A graphical representation of a 'perfect' competitive market consisting of $n = 16$ consumers. (b) The same $n = 16$ consumer market subjected to trade restrictions.

Even from this simple representation, we can derive several important structural differences between the two markets. First, from a network perspective the perfectly competitive market in figure 5.1(*a*) would be considered a *complete graph*, (often written as K_n, in this case $n = 16$), in the sense that every pair of vertices is connected by an edge. As a consequence, we can also state that the perfect market is a *regular graph*, such that each vertex has the same number of neighbours, i.e., each vertex degree $k = n - 1$. On the other hand, the restricted market[1] shown in figure 5.1(*b*) is sparse of edges between consumers, and displays heterogeneity in terms of the number of neighbours connected to each consumer (see Chapter IV for methods

[1] The network displayed in figure 5.1(*b*) was drawn from a *random graph* model [66], where every pair of vertices is connected by a edge with a probability $0 < p < 1$.

www.manaraa.com

of measuring degree heterogeneity). The result is a wide range of vertex degrees, between $1 < k < n-1$[1].

To begin to understand how these differences in market structure affect price formation, we need an economic model of consumer behaviour to apply onto the network model. While the field of network science is largely focused on developing models that are purely structural (see Chapter III for examples), one model called the *graphical exchange model* combines both network effects with economic logic of buyers and sellers in a market. Developed over a series of papers by Kakade *et al*. [see 85, 86, 84, 132], the graphical exchange model sets all standard economic factors; endowments, utilities, and so on, to be, both uniform and symmetrical across all consumers. This design, ensures that any resulting price variation at equilibrium, can be ascribed solely to the underlying structure of the market under study [132]. As a result, consumers may have to pay dramatically different prices for the same commodity, based solely on their privileged position in the network. This chapter focuses on understanding how specifically, the topology of a network affects the outcome of trade using this model. In the next section we give some background on a standard economic model that is used throughout the literature, and from which, the graphical exchange model is based.

5.2 The Classical Fisher Market Model

In this section we briefly explain the classical Fisher market model [53], from which the graphical exchange model [132] is a specialised derivation. The Fisher model describes a competitive commodity market in which there are m buyers of n goods[2]. The aim of each buyer is purchase g_j units of each good j at a a price p_j, that maximises the buyer's utility u_i. Each buyer i is given a cash endowment e_i from which to make these purchases. The main assumption[3] of the model is that the utility function of each buyer is *linear* in the amount of goods purchased [84], i.e., the more of the good j, buyer i can obtain, the more utility (or 'value') they will gain from the purchase. If u_{ij} denotes the utility derived by i on obtaining a single unit of good j, and buyer i purchases x_{ij} amount of good j, then the utility i derives is given by,

$$\sum_j u_{ij} x_{ij}. \tag{5.1}$$

[1] A consumer with $k = 0$ is not able to trade with anyone, and is therefore not considered part of the market.
[2] The goods are being sold by a number of sellers presumably, but they are not explicitly named in the model [53].
[3] Along with other standard economic modelling assumptions: strictly monotone/weakly convex preferences, a constant marginal rate of substitution, and perfectly substitutable goods. For details, see [59].

www.manaraa.com

In this market, a set of prices $\mathbf{p} = \{p_1, p_2, \ldots, p_j\}$ will reach equilibrium if the following two conditions hold:

- The market clears (all goods are sold), so that supply equals demand, i.e., $\sum_i x_{ij} = g_j \quad \forall j$.
- The total cost of the goods purchased by each buyer i is not more that the buyer's cash endowment i.e., $x_{ij} \cdot \mathbf{p} \leq e_i \quad \forall j$.

Given these assumptions, Arrow-Debreu [8] proved that an equilibrium will always exists for some set of prices $\mathbf{p}$ if each good j has a buyer which derives non-zero utility for that good, i.e., $u_{ij} > 0$ for some i. The problem has been formulated more formally [see 37] as follows,

$$\begin{aligned} \text{maximise} \quad & \sum_{i=1}^{m} e_i \log u_i & \\ \text{subject to} \quad & u_i = \sum_{j=1}^{n} u_{ij} x_{ij} \leq 1 & \forall i \\ & \sum_{i=1}^{m} x_{ij} \leq 1 & \forall j \\ & x_{ij} \geq 0 & \forall i, \forall j. \end{aligned} \tag{5.2}$$

Note, that the existence of equilibria does not mean that equilibrium is always achieved. A mechanism is required to reach convergence, such as auctioning or bargaining for example [9]. The idea suggested in Fisher [53], is a Walrasian-esque (after Léon Walras) "tatonnment" process, where each buyer submits a list of purchases x_{ij}, for every possible price p_j of each good j. A third-party then calculates which price p_j will satisfice the total demand across all buyers. Clearly, this is not a particularly realistic mechanism, given that there is no known market that works in this manner, but it is at least believable, in the sense that all the information required to calculate equilibrium prices (the buyers own utility u_{ij} and endowment e_i) are known to the people making the decisions. The actual calculation of prices in a Fisher market model has proven difficult due its computational complexity, and is still an active area of research [9].

The important takeaway from this model is that the market has no effect on wealth across the market. That is to say, each buyer starts with endowment e_i and ends with e_i's worth of a selection of j goods, depending on their preferences. In this way, the process can be thought of as a kind of matching algorithm, matching interested buyers to sellers. What is not shown in the model is that this is a function of the market's structure. The hidden assumption in the Fisher model is that it is operating in a perfectly competitive market, like the example shown in

www.manaraa.com

figure 5.1(*a*). What is not known, is how does relaxing this assumption, and running the model on a restricted market, affect both the distribution of wealth and prices.

In the next section, we introduce the *graphical exchange* model, a specialisation of the classical Fisher model[1] [53], that was recently developed in a series of papers by Kakade *et al*. [see 85, 86, 84, 132]. The authors note that from an economics point of view, the graphical exchange model offers no advancement over the Fisher model, but does allow for the underlying structure of the market to influence the properties of equilibrium prices [84]. As such, this model is key to allowing us to answer the research question posed above.

5.3 A Graphical Exchange Model of Trade

In this section we introduce the graphical exchange model [following 84, 132]. The model describes a competitive market, consisting of n consumers, divided into an equal number of i buyers and j sellers. For the sake of simplicity, the number of goods sold in this market is limited to one, herein called *wheat* (without loss of generalisation to other commodities) and that all wheat sold is homogeneous, infinitely divisible, and perfectly substitutable[2]. The aim of each buyer is to purchase g_j units of wheat from any seller (or from several sellers), that maximises the buyer's utility u_i. Each buyer i is given a cash endowment e_i from which to make these purchases. The same linear utility assumption from the Fisher model applies, in that the more wheat buyer i can obtain, the more utility (or 'value') they will gain from the purchase. In this market, a set of prices $\mathbf{p} = \{p_1, p_2, \ldots, p_j\}$ will reach equilibrium if the following two assumptions hold:

- The market clears (all wheat is sold), so that supply equals demand.
- The total cost of the wheat purchased by each buyer i is not more that the buyer's cash endowment.

Given these assumptions an equilibrium will always exists for some set of prices [84]. Where the graphical exchange model diverges from the Fisher model, is that each seller j is explicitly named and is given an endowment g_j units of wheat to be sold at a *locally determined* price p_j. That is to say, the model permits and realises prices that can diverge between sellers. These price differences for identical goods are assumed to arise due to the network structure of the underlying economic interactions [85].

[1] Graphical Arrow-Debreu models have also been developed [see 86].
[2] Note that these are standard simplifying assumptions in many economic models.

www.manaraa.com

5.3.1 Representing the market as a network

In a network setting, we can represent the market as a directed bipartite graph, $G = (U,V,E)$, comprising two sets of vertices, U consisting of $\{1,2,\ldots,j\}$ sellers, and V consisting of $\{1,2,\ldots,i\}$ buyers, as well as a set of edges E connecting buyers to sellers across sets. Each edge between buyer i and seller j denotes where free trade is allowed between the two parties, while the absence of an edge denotes a restriction on trade. Naturally then, buyers are only allowed to purchase wheat from a seller with whom they are connected, and *vice versa*. We can record trade restrictions in the graph's adjacency matrix $\mathbf{A}$, whose elements $a_{ij} = 1$ where trade is permitted and $a_{ij} = 0$ where trade is not. In a bipartite system, trade is not possible with others in the same set (i.e., buyers to buyers), so that all permitted edges in E are between sets U and V.

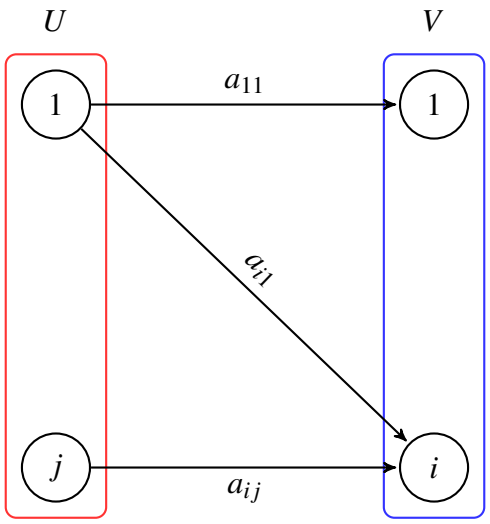


Fig. 5.2 The bipartite graph $G = (U,V,E)$, where U is a set of nodes representing sellers and V is a set representing buyers. Edges from U to V exists if and only if trade is permitted, where $a_{ij} = 1$. Note that the edge between buyer 1 and seller j is not present so $a_{1j} = 0$.

Given the network market model outlined above, we want to be able calculate the set of prices $\mathbf{p}$ that clears the market, both locally and globally, such that, all wheat is sold, and all cash endowments are used both at the individual level and across the market as a whole. In the classic Fisher model, assuming only one good, we could simply define a total demand vector as $\mathbf{d} = \sum_i x_i$, where x_i is the total amount of wheat purchased by buyer i, and a total supply vector as $\mathbf{g} = \sum_j g_j$. Thus, the market clears if $\mathbf{d} = \mathbf{g}$. If the graphical exchange model is run on a complete network (where every pair of vertices is connected by a edge), then these notations for market equilibria would suffice. However, if the graph constrains trade between

www.manaraa.com

individual buyers and sellers, the assumptions of the Fisher model do not hold [85]. With this in mind, Kakade *et al*. [85] developed the notion of *local market clearance*, in which equilibria is achieved if and only if all wheat is sold within the *neighbourhood* of each seller[1]. Here, the local demand vector is defined as $\mathbf{d}_j = \sum_{i \in U(j)} x_{ij}$, so that the local clearance condition for each j becomes $\mathbf{d}_j = g_j$. The market clears globally when $\mathbf{d}_j = g_j$ holds for $\forall j$. Kakade *et al*. [85] argue that this notion of *local clearance* more clearly expresses the economic logic of each seller in the network, compared to the global approach of the Fisher model. For example, the United States government is very much interested in balancing the production of American goods with the current (and predicted) worldwide demand for those goods, whereas the balancing of worldwide production and worldwide demand would only be considered in terms of its secondary effects on American interests.

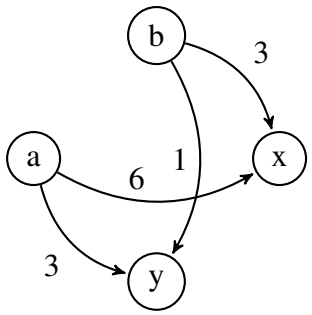
5.3.2 Calculating equilibrium prices using max-flow/min-cut theorem

To calculate the local supply and demand for wheat, we employ the *max-flow/min-cut* theorem to calculate the 'flow' of wheat across the network [30]. The idea, is that the flow of wheat through the system is constrained, first by the network structure, but also by the limited 'capacities' c of each consumer to either supply or demand. Supply here, is driven by prices, so that where there are higher prices, sellers will want to supply more of the good. In theory, demand is infinitely elastic, which is to say the quantity of wheat purchased would increase without limit as prices tend to zero. However, in practice, demand is limited by the total cash endowment of all buyers in the market. To induce the flow across the graph, a 'source' vertex s and a 'sink' vertex t are added to the network to create a flow-graph F (these are engineering metaphors). The maximum flow $|f|$ of the flow-graph tells us how much wheat can flow from the source—to the sink without exceeding the capacities of any of the consumers involved [30]. A way of solving this problem is to find the bottleneck in the network's structure that is constricting the flow of wheat. A bottleneck is typically identified by a 'cut' or partition of the network, that will split the flow-graph into two disjointed sets of vertices $(P, \bar{P})$, denoted by a list of edges that would need to be removed in order to split the flow-graph into P and $\bar{P}$ [30]. The minimum cut of the flow-graph is where the flow along edges *between* P and $\bar{P}$ are at their maximum capacity, and is denoted by $c(P, \bar{P})$. The max-flow/min-cut theorem states that the maximum flow $|f|$, between the source vertex s and the sink vertex t is equal to the sum capacities of the minimum cut $c(P, \bar{P})$ [30]. This is important, as in complex cases there could be multiple

[1] Neighbourhood being defined here as a sub-graph of the network consisting of all vertices adjacent to a given vertex.

www.manaraa.com

different flow configurations, and going through them one by one to find the maximum flow can be prohibitive in terms of the time complexity of the problem [85].



$$\mathbf{W} = \begin{bmatrix} 0 & 0 & 6 & 3 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Fig. 5.3 (Left) A simple bipartite trade network, with sellers a and b trading with buyers c and d subject to constraints. (Right) the corresponding weighted adjacency matrix $\mathbf{W}$, where the element w_{ij} indicates the capacity of the edge from i to j.

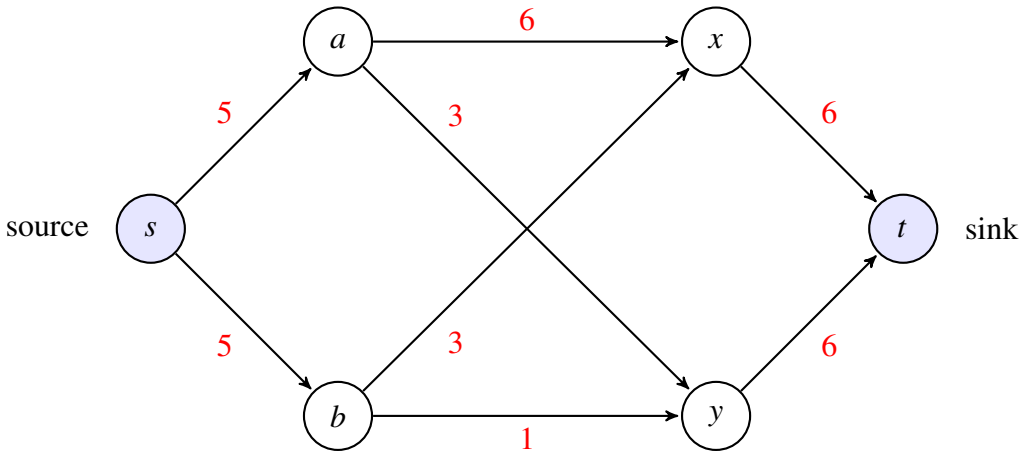


Fig. 5.4 A flow-graph of the example network shown in Figure 5.3 above. Note the source vertex s is attached to a and b with an asking price of 5, and vertices x and y are attached to the sink vertex t, each with an endowment of 6.

In our exchange model (see Appendix B1, Listing B.1) the max-flow and min-cut are calculated using the `graphmaxflow` function in Matlab's bioinformatics toolbox [98], which is an implementation of the Goldberg "preflow-push" algorithm [69]. To illustrate this method with an example, we work through a typical max-flow problem below. In figure 5.3 (*left*), we create a simple bipartite trade network, with two sellers, a and b, and two buyers x and y. Buyers can trade with each seller, but not with each other, and *vice versa*. Imagine that these trade relationships are subject to some constraints, indicated by the edge weights in the adjacency matrix (see figure 5.3 (*right*)). To compute the maximum flow across this graph, we

www.manaraa.com

first need to build an augmented flow graph of the market. In figure 5.4, we add a source vertex s, which is attached to sellers a and b with an asking price of 5. Buyers x and y are each given an endowment of 6 units of cash with which to make their purchases, and attached to a new sink vertex t. If we assume a and b both have one unit of a good to sell, the buyers have more than enough cash to purchase all goods with money left over (the total cost of goods offered is $(1\times 5)+(1\times 5)=10$, and the total cash endowment is $6+6=12$). However, the trade of goods in this network is limited by a set of constraints between buyers and sellers. These constraints are represented by the maximum capacity c of each edge, denoted in figure 5.4 by the edge labels in red. So for example, we can say the edge between vertices a and x, (a,x) has a maximum capacity $c(a,x)=6$. In figure 5.5, we induce the flow through the flow-graph given these constraints. Now, the second label in blue denotes the current flow f being pushed through the network. We can see that while the maximum capacity of (a,x), $c(a,x)=6$, the flow is only $f(a,x)=3$. In this example, the maximum flow of the network can be found from the sum of the current flow to the sink vertex t, from vertices x and y, where $f(x,t)=6$ and $f(y,t)=3$, giving a maximum flow $|f|=6+3=9$. The min-cut of the flow-graph divides the network along the edges $(s,a),(b,x)$ and (b,y) so that $P=\{s,b\}$ and $\bar{P}=\{a,x,y,t\}$. Note that true to the max-flow/min-cut theorem, the capacity c of the edges between P and $\bar{P}$ also sum to equal the maximum flow of the network $c(P,\bar{P})=5+3+1=9=|f|$.

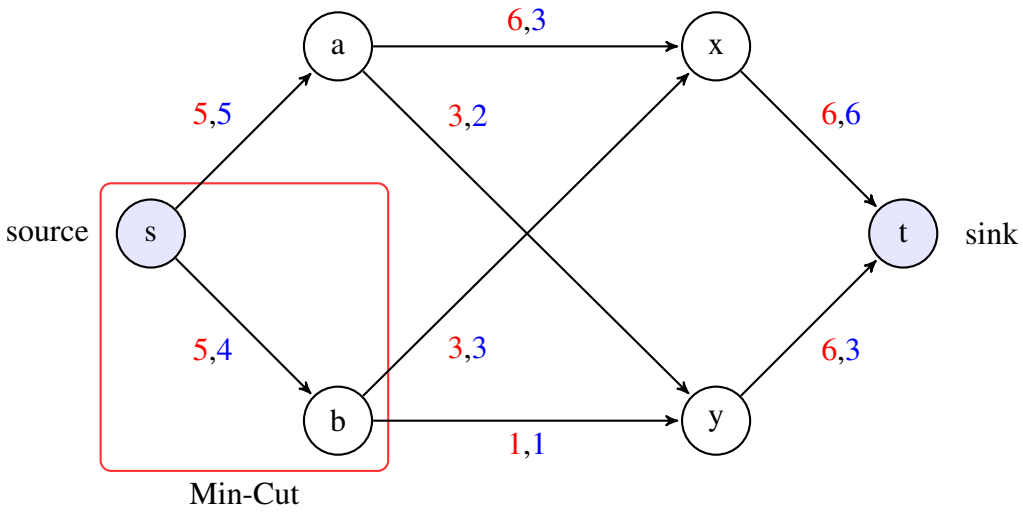


Fig. 5.5 An example max-flow problem. The edge label in red denotes an edge's capacity c, the edge label in blue denotes the current flow f. The set $P=\{s,b\}$ (outlined in red) is the min-cut of the graph.

www.manaraa.com

5.3.3 Application to the graphical exchange model

To apply the max-flow/min-cut theorem to the graphical exchange model, we first need to build an augmented flow graph of the market. Following [86], we create a flow-graph from the network $G = (U, V, E)$, shown in figure 5.2, as follows; we first assign a capacity of infinity to all the edges from U to V in G. That is to say there is unlimited free-trade permitted between any pair of consumers that are connected by an edge. Next, a source vertex s is introduced and connected by directed edges from s, to each seller $j \in U$, with a capacity of p_j representing the asking price at each vertex. On the other side of the graph, a sink vertex t is introduced and also connected by directed edges from each buyer $i \in V$ to t with a capacity of e_i representing the buyer's cash endowment. This makes sure that each buyer cannot spend more money than is available to them. The resulting flow-graph, shown in figure 5.6, is a function of the current set of prices $\mathbf{p}$ and can be denoted as $F(\mathbf{p})$.

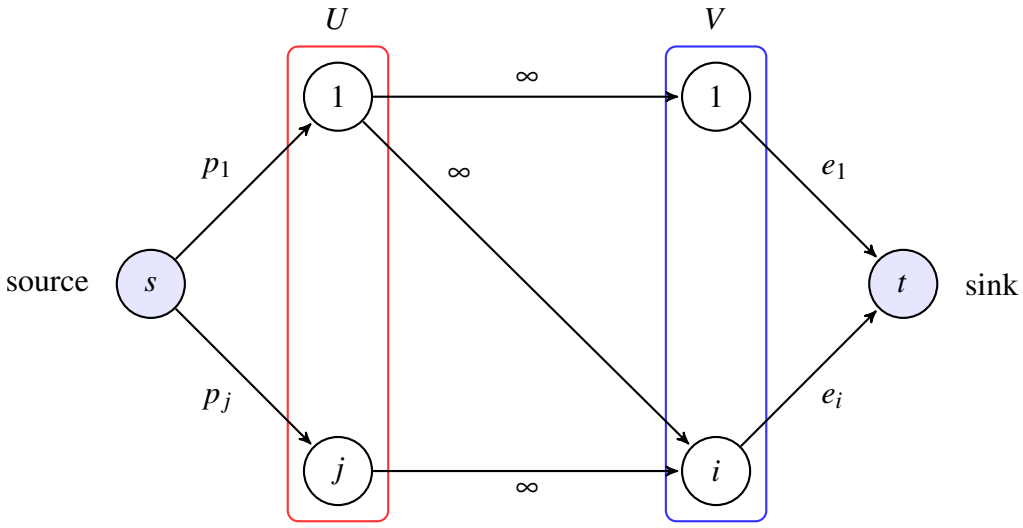


Fig. 5.6 An augmented flow-graph of the graphical exchange model, shown in Figure 5.2.

This augmented flow-graph is used to find a set of prices $\mathbf{p}$, that will clear the market both locally and globally. We use the max-flow/min-cut method outlined above as one part of a primal-dual algorithm, in a similar fashion to the market equilibrium algorithm proposed by Devanur et al. [37]. The idea here, is to start raising prices across the flow-graph until a subset of sellers $R \in U$, and neighbouring buyers $N(R) \in V$, find a price that leads to a *local* optimal set of trades. We can determine when a set of trade become optimal, as the min-cut will shift to be around both the source vertex s and the optimal subset $(R, N(R))$. We then freeze any subset of optimal trades $(R, N(R))$ and only raises prices in the remaining active sub-graph $(U - R, V - N(R))$. This is to ensure what Devanur et al. [37] call the 'invariant

www.manaraa.com

condition'—that the min-cut must always include the source vertex, such that $(s, U \cup V \cup t)$, meaning all goods can be sold at current prices $\mathbf{p}$. This process continues, finding and freezing more subsets in $F(\mathbf{p})$ as prices keep rising. To ensure that a min-cut can always be found, we first set prices $\mathbf{p} = 1/\max(k_j)$, where k_j is the degree of each j seller. Prices are then raised systematically by multiplying $\mathbf{p}$ by $x = 1$, then $x = 1.1$, $x = 1.2$ and so on. We clear the market globally by running the algorithm until all subsets of buyers and sellers have found a locally optimal trade. This will be equivalent to the condition that the min-cut is now around the sink vertex, i.e., $(s \cup U \cup V, t)$.

5.4 The Determinants of Local Price Variation

Now that we have a method of finding equilibrium prices using the graphical exchange model, we want to look into the determinants of price variation in networked markets at a local level. In particular, we are interested in whether specific buyers or sellers are able to command favourable prices due to their position in the network. For example, microeconomic theory tells us that in a given situation where demand is greater than supply, a shortage occurs, leading to a higher equilibrium price [20]. We might presume, therefore, that a seller with many interested trade partners (i.e., with a high degree k) will be able to command a high price for their goods.

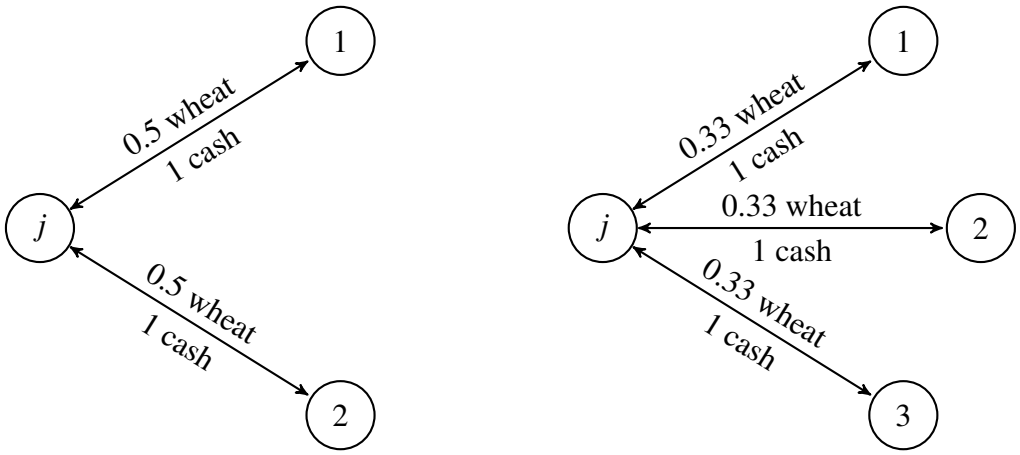


Fig. 5.7 (Left) the local neighbourhood of seller j, $N(j)$ with two interested buyers. At equilibrium seller j trades 0.5 units of wheat to each interested buyer in exchange for 1 unit of cash each. The market clearing price is 2 units of cash per unit of wheat sold. (Right) the local neighbourhood of seller j with three interested buyers. At equilibrium seller j trades 0.33 units of wheat to each interested buyer, again for 1 unit of cash each. The market clearing price in this case is 3 units of cash per unit of wheat sold.

www.manaraa.com

To illustrate this point, we consider only the local neighbourhood $N(j)$ of seller j (i.e., the seller and its trade partners) in figure 5.7. First, where j has two interested buyers, and second with three. Recall that in the graphical market model each seller is endowed with only 1 (divisible) unit of wheat and each buyer is endowed with 1 unit of cash. Ideally, each buyer would like to trade their endowment for as much wheat as possible. However, there is too little wheat on offer to satisfy all demand. As such, the interested buyers will have to compete by offering increasingly more cash per unit of wheat, until j has sold all of their wheat endowment. In the first case, at equilibrium, seller j would trade half of their endowment to buyer 1 and half to buyer 2. The resulting price (expressed as a unit/price ratio) would be 2 units of cash per unit of wheat sold (herein we express prices simply in terms of a dollar value, so a \$2 wheat price is equivalent to trading 2 units of cash per unit of wheat). In the latter case, we see that seller j would trade $1/3^{\text{rd}}$ of their wheat to each interested buyer, at a price of \$3 per unit. We can see from this example that, within a seller's local neighbourhood, the maximum price they can charge is bounded by the number of interested buyers. More generally, if the graph G consists of only the immediate neighbourhood of seller j (that is j and its k buyers), then the equilibrium price in G is equal to the degree k, since all buyers are forced to buy from seller j [86].

While it is understood that prices in the graphical economy model are bounded above by the seller degree k (which is the most basic measure of local network structure), Kakade et al. [86] find that k, is in fact, a poor predictor of individual seller price. This is because sellers face competition, and competition drives down prices. Graphically, we can measure local competition using the average nearest neighbours degree k_{nn}, which takes the average degree of the k importers in the neighbourhood of seller j. In this setting, larger values of k_{nn} indicate that seller j is competing with many other sellers in the market. We can represent the varying levels of competition k_{nn}, and seller degree k, in the local market structure using idealised examples from microeconomics. Following Fernandez *et al.* [51], we display these recurrent structures in figure 5.8 using graphical representations, often referred to as network *motifs*.

The first network motif in figure 5.8(a) shows a localised monopoly, where the highly connected seller j is the sole supplier to many interested buyers. Since, all buyers in j's neighbourhood are dependent on j for their supply, they have no option than to pay full price. In the graphical exchange model, the resulting equilibrium price in market structured as this motif would be \$5 (see table 5.1). The monopsony motif in figure 5.8(c), is simply the opposite of a monopoly. Here, many sellers are competing with each other to trade with a single buyer. The extreme competition in this structure drives down prices, such that a seller can only

www.manaraa.com

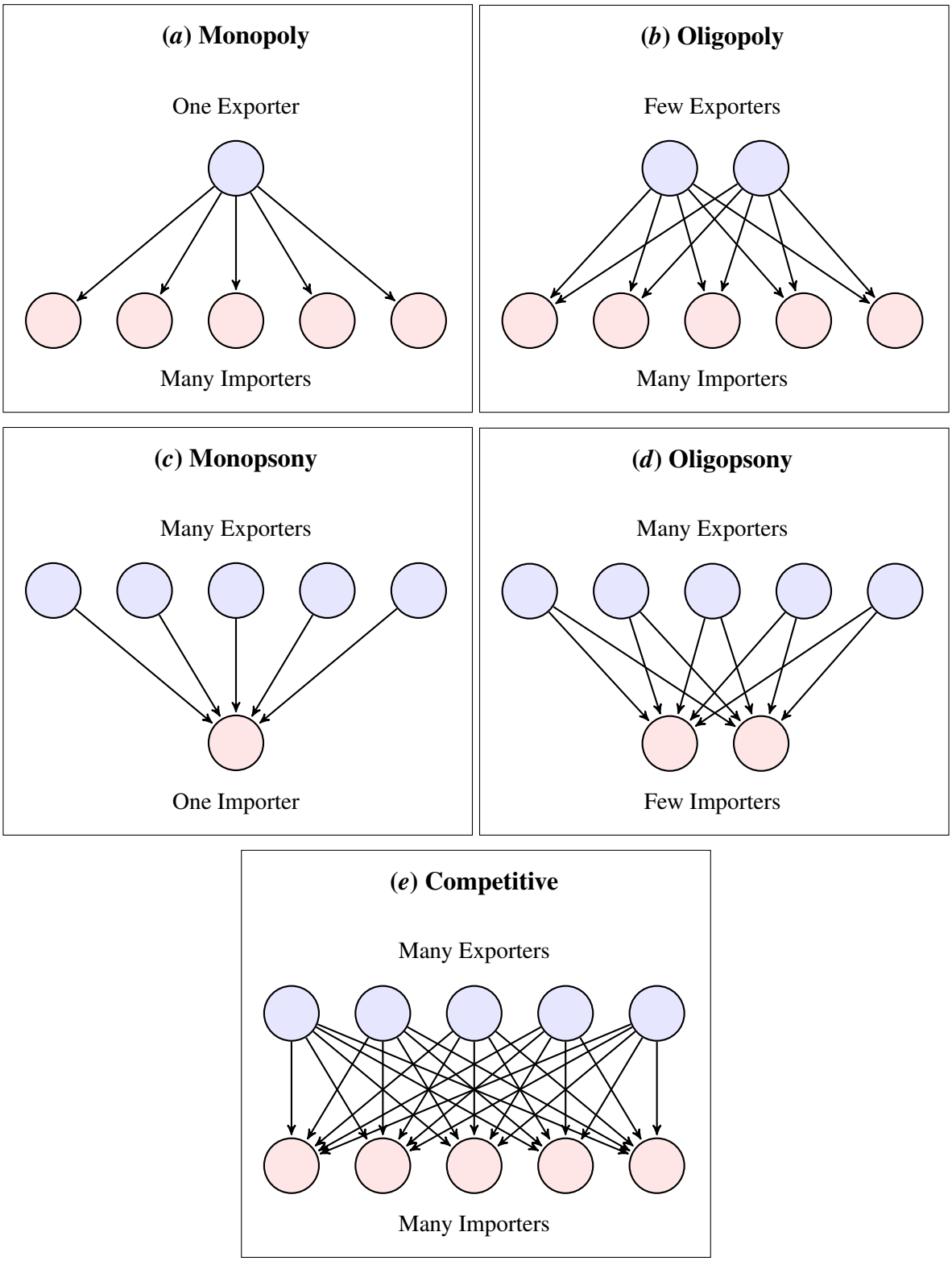


Fig. 5.8 Local network motifs developed by Fernandez *et al*. [51]. See text for description.

www.manaraa.com

command \$0.2. The prices in the monopsony network motif are 25 times less than the price sellers can charge in a monopolistic market. Given the network structure of these motifs, we can assume that any observations of highly connected buyers or sellers in market data may indicate the presence of either a monopoly or monopsony.

The oligopoly network motif shown in figure 5.8(*b*), is where there is more than one seller, but few compared to the number of buyers. Since the market comprises more than one seller, an element of competition is introduced. We can see from our example that the sellers each have the same degree $k = 5$, as in the monopolistic example, however, now that there are two sellers, the average nearest neighbours degree $k_{nn} = 2$. This small amount of competition halves the price these sellers can command in the graphical economy model. Again, the oligopsony motif shown in figure 5.8(*d*) is simply the opposite of an oligopoly. Here, the low number of buyers in relation to sellers drives down prices markedly. Finally we have a competitive network motif in figure 5.8(*e*), where there is an equal number of buyers and sellers. As such, each seller in this market has the same degree k and average nearest neighbours degree k_{nn}. In effect, the network structure in this configuration has no influence on prices in the graphical economy model. In microeconomic theory, a perfectly competitive market is an ideal case where neither buyers or sellers have power to influence price over the rest or the market [51]. Rather, aggregate market supply and demand determines prices.

From the values in table 5.1 below we can clearly see that k is not the sole determinant of prices, given that the first three motifs (Monopoly, Oligopoly, and Competitive) have the same degree but prices ranging from \$5 to \$1. Rather, in these idealised examples, it is the ratio of k to k_{nn} that determines prices. Unfortunately, real-world markets are rarely organised in this way. For example, whilst seller j could form a monopolistic subgraph in relation to the majority of its neighbours $N(j)$, some connected buyer $i \in N(j)$ may also be connected to a largely competitive subgraph in a different part of the market. The competition in the highly connected subgraph could therefore, *indirectly* influence prices in the adjacent monopolistic subgraph, with i operating as of form of arbitrage agent. As such, we cannot determine the exact price commanded by a seller, based on knowing only k and k_{nn}. However, we can determine the upper and lower bounds these values may take. This is achieved by measuring both seller degree as a function of price $k(p)$ and the average nearest neighbours degree as a function of prices $k_{nn}(p)$, for all values of k and k_{nn}.

www.manaraa.com

Table 5.1 Local network characteristic of varying market structures. Values denote: exporter degree k, average nearest neighbour degree k_{nn}, and local equilibrium price p determined by the graphical exchange model.

Local Market Structure	k	k_{nn}	p
Monopoly	5	1	\$5.0
Oligopoly	5	2	\$2.5
Competitive	5	5	\$1.0
Oligopsony	2	5	\$0.4
Monopsony	1	5	\$0.2

5.5 Main Results

5.5.1 Theoretical Bounds on Price Variation

To test these theories on larger scale networks, we take a numerical approach to determining the theoretical upper and lower bounds of price variation across a market. A set of bipartite graphs were generated in which every seller has the same degree k and the same average nearest neighbour degree k_{nn}. Measurements of equilibrium prices were then taken for all values of k and k_{nn}. Given that prices in these theoretical markets are only relatable to each other, the observed price variation p, measured as a percentage change from \$1 (i.e., the value expected in a complete/competitive market), has been normalised on a scale of -1 to 1. So that where $p = 1$, prices are 100 times greater than in a competitive market and where $p = -1$, they are 100 times less.. Note that the values for the degree k and the average nearest neighbour degree k_{nn}, are derived from the adjacency matrix of a graph G, i.e., the original network representing the ability of buyers and sellers to trade, rather than the flow-graph $F(\mathbf{p})$, which has been augmented with the addition of sink and source vertices. This is to ensure that the effect on price variation is determined by the restrictions on trade—which are encoded in G.

The upper plot in figure 5.9 shows that our results are in line with the theory that put forward by Kakade et al. [86], that the upper bound ($p^{\max}$) of the variation in prices expected by a seller is determined solely by the seller's degree k.. Note that the large gaps in the top half of the plot is caused by the gradients of k and k_{nn}. We can also see that the lower bound ($p^{\min}$) is determined by k/n, where n is the number of sellers in the market. This value represents the price seller j would expect if they had to compete with every other seller in the market. At the lower bound, where $k = n$, the price variation $p = 0$ as the graph is complete and functionally equivalent to the classical Fisher market model. That is to say there is effectively no network

www.manaraa.com

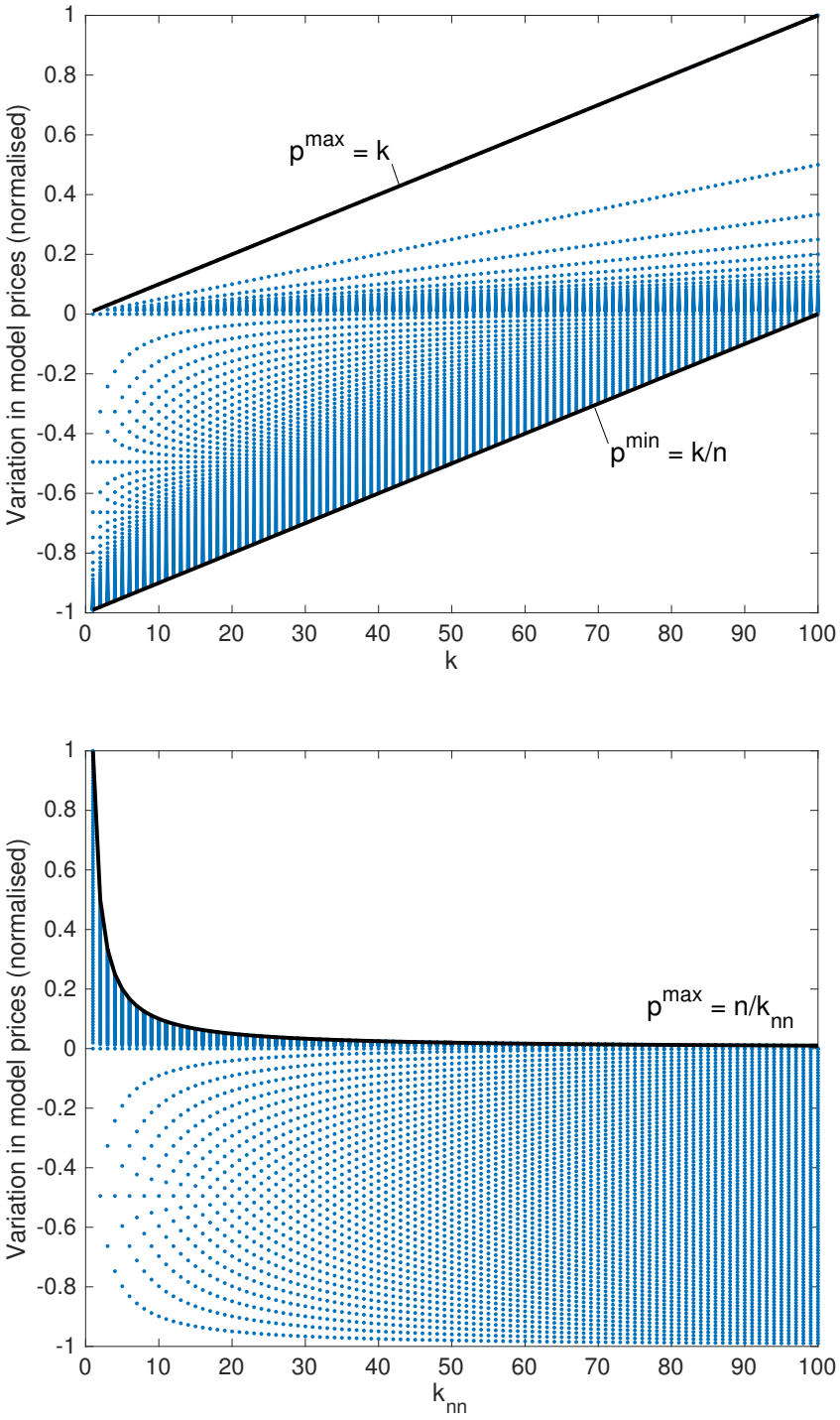


Fig. 5.9 (Above) variation in prices p as a function of degree k. (Below) price variation as a function of the average nearest neighbour degree k_{nn}.

www.manaraa.com

structure to affect prices. As such, we can say that increasing a seller's degree k (i.e., their access to markets) monotonically increases their prospects of obtaining better prices.

The lower plot in figure 5.9 shows price variation as a function of the average nearest neighbour degree k_{nn}. We can see that prices are bounded above in a power-law fashion, where $p = n \times {k_{nn}}^{-1}$. We can interpret k_{nn} as the amount of competition each seller faces for each trade. Where $k_{nn} = 1$ and $k = n$, a seller has a complete monopoly over trade, such that they can command 100 times greater prices than in a competitive market. However, as soon as competition is introduced, such that $k_{nn} = 2$, the upper bound on price variation ($p^{\max}$) is halved. Where $k_{nn} = 3$, $p^{\max}$ is cut to one third, and so on. As such, we can say that competition (in the form of a seller's second nearest neighbours) rapidly eats into a seller's ability to command higher prices. Note that $p^{\min}$ has no lower bound in relation to k_{nn}.

In figure 5.10 we plot a heatmap of seller degree k over the average nearest neighbour degree k_{nn}, with the colour denoting the resulting variation in prices p. Unsurprisingly, given the results in figure 5.9, the distribution of price variation when k is plotted against k_{nn} is not symmetrical between positive and negative changes in price. Starting from along the vertical, price variation decreases monotonically with degree k until k and k_{nn} achieve parity. There is no price variation along the diagonal as this area represents a *locally* complete sub-graph of the market. Below the diagonal, price variation rapidly increase as k_{nn} decays. This leaves a small zone of local monopolistic prices along the horizontal. Given a market structure in which all value of k and k_{nn} are equally likely (i.e., a random graph), the resulting variation in prices would be somewhat biased towards lower prices. Simulations show an expected mean price across a random graph 0.22 times less than in a complete market.

5.5.2 Application to Real-World Data

To provide some context to the theoretical results outlined above, we apply the graphical exchange model to trade data representing the real-world international wheat market. Much like in the previous chapter (Chapter IV), we use data from the UN Comtrade database [141] detailing import/export transactions between countries in the year 2015. These data include, for each pair of countries, the total volume of wheat (measured in current U.S. dollars worth) trade in each year[1]. From these data we extract the discrete network structure of the market using the matlab function shown in Appendix A1, resulting in a $n \times n$ adjacency matrix (where $n = 79$) much like the bipartite graph outlined in section 5.3.1. So here, the buyers and sellers

[1] HS-02 classification system was used to select the product code; 1001 denoting trade in wheat

www.manaraa.com

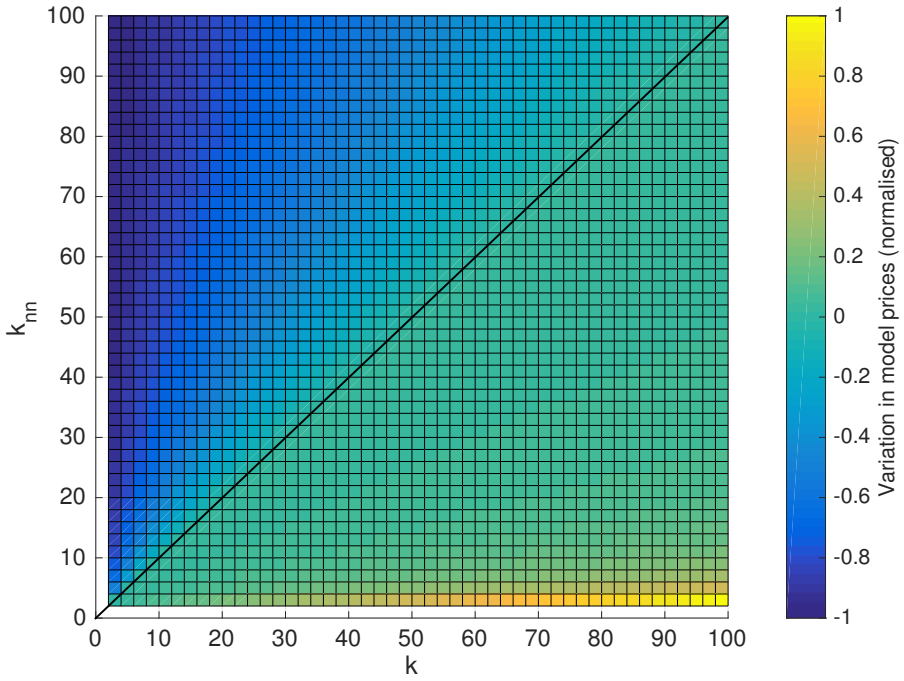


Fig. 5.10 Heatmap of price variation, showing exporter degree k against the average nearest neighbour degree k_{nn}.

in the network now represent the importing and exporting countries, and the edges represent where trade was possible between those countries.

In their analysis of aggregated trade data, Kakade et al. [86] randomly assign roles to each country, as either a buyer (importer) or a seller (exporter). However, at the individual commodity level, we would argue that the directed and bipartite nature of this trade network should be preserved, as it is an important factor in the determinants of price formation. For instance, the majority of countries that export wheat, do not import significant amounts in return, and *vice versa*. As such, the way exporters are connected to importers is different to that of the importers [52]. More concretely, countries with these different roles have been shown to have marked differences in terms of their local topological characteristics [52]. Therefore, we maintain the specific role each country takes in the data, so that the structural asymmetries presented, are a direct result of the directed-bipartite nature of the network.

Figure 5.11 shows international wheat exporters ranked in descending order. The basis of these rankings is the local equilibrium price p derived from using the graphical economy model applied to the UN Comtrade data [141]. Given the specification of the model, prices p are in

www.manaraa.com

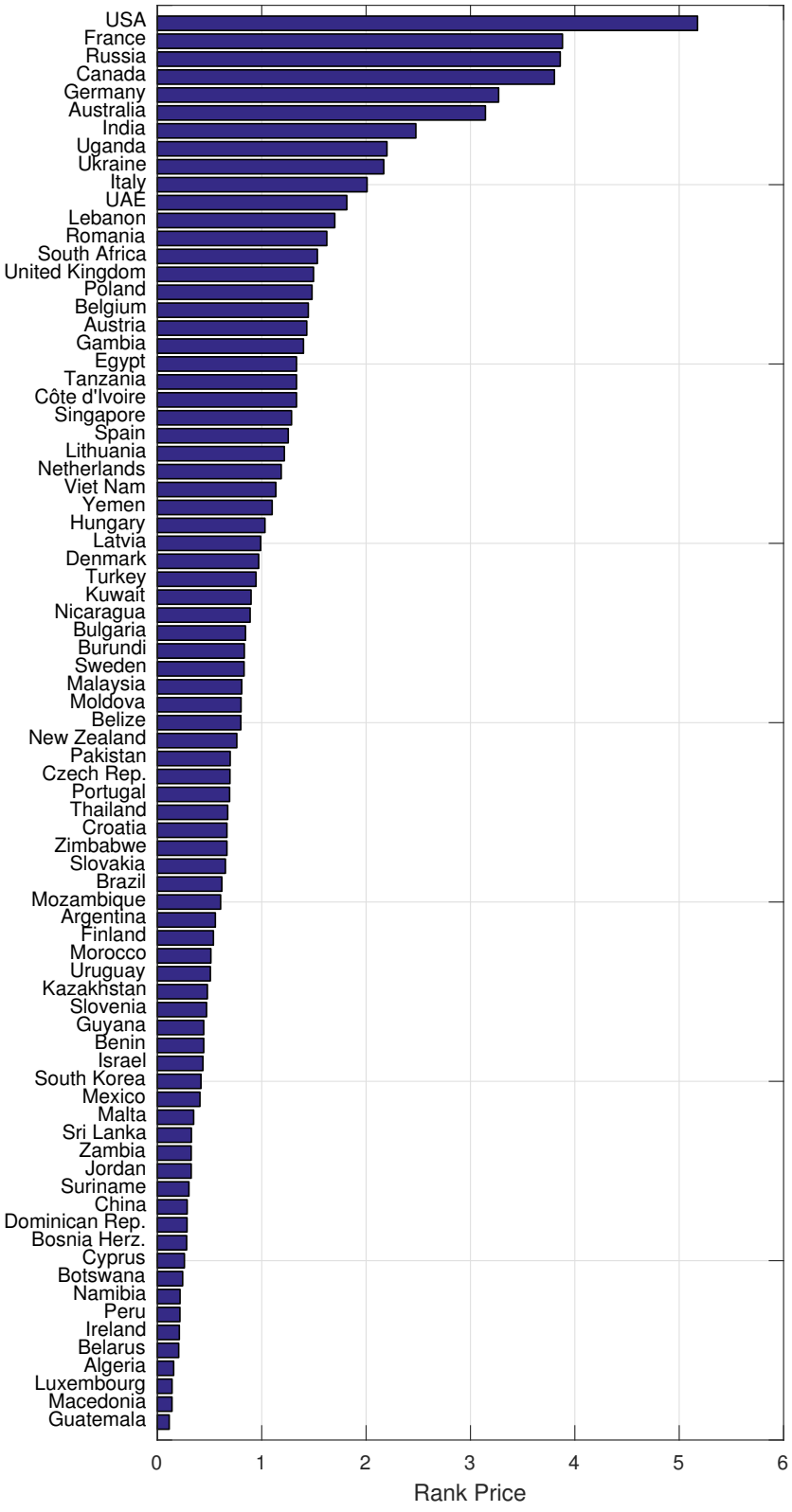


Fig. 5.11 Ranked wheat exporters based on model prices (network data from UN Comtrade database [141]).

www.manaraa.com

effect 'shadow prices'[1] as they only take into account the network structure of trade, and not the overall supply and demand conditions. Model prices are therefore not directly comparable to real-world market prices.

From the rankings, we can immediately see a high degree of price variation due to the underlying network structure of the market. Prices across the network range from 0.11 to 5.2. While the majority of exporters (50/79) can expect $p < 1$, the top ten can expect $p \geq 2$. The exporters with best market position command prices of 5.2 (USA), 3.9 (France), 3.9 (Russia), 3.8 (Canada), and 3.2 (Germany). We can also see that the distribution of prices is skewed towards those at the top. The USA in particular stands out, having a 33% higher price than its next-nearest competitor.

Table 5.2 Top ranked wheat exporters ranked on model prices (network data from [141]). Values denote: local equilibrium price p determined by the graphical economy model, the total export value (in US$), and the total export volume (in Millions of metric tonnes).

Top Ranked Wheat Exporters				
Country	Model Rank	Model Price	Export Value (US$)	Export Weight (Mmt)
USA	1	5.17	304 (1)	252 (1)
France	2	3.88	182 (3)	159 (2)
Russia	3	3.86	3 (17)	3 (16)
Canada	4	3.80	197 (2)	983 (3)
Germany	5	3.27	59 (5)	43 (5)
Australia	6	3.14	1.8 (22)	2.4 (18)
India	7	2.47	0.3 (30)	0.3 (29)
UAE	8	2.20	0.09 (51)	0.04 (47)
Ukraine	9	2.17	0.15 (24)	0.15 (23)
Italy	10	2.01	3 (18)	1.7 (22)

In table 5.2, we compare this price ranking to real-world export values and volumes, again taken from the UN Comtrade database. The figure in parenthesis next to each value is the ranking of that exporter based on either the total value (in US$) or the total volume (in Millions of metric tonnes) of wheat traded. We can see that four of the top five exporters follow similar rankings to the model prices, with Argentina being the missing exporter from this list (ranked 4th by both value and volume). However, Russia performs far better in the model than in the other rankings. We can also see that, as with the model prices, the U.S. outperforms all other countries in both the value and volume of wheat exported.

[1] Shadow prices are values assigned to unknowable or difficult to calculate costs [39].

www.manaraa.com

Although somewhat counter-intuitive, we should not necessarily expect model prices to correlate with these rankings. First, the model specification rules out externalities, such as government interventions, that might affect export values and weights. This allows for a measure of exporter performance based solely on network position, but naturally, limits the scope of the model to capture all real-world eventualities. Second, while the export value and export weight can be considered a measure of exporter performance, they measure each exporter against the global market. The model prices, on the other hand, can be considered a measure of how well a country performs within its own *local neighbourhood*. That is to say, some of these exporters do not necessarily compete directly with one another. Thus comparing all exporters globally might not accurately reflect their model ranking. For example, Russia, and Ukraine, along with Kazakhstan to a lesser extent (Model Rank 19), dominate exports from the Black Sea ports. Countries in Central Asia, that are dependant on 'Black Sea' wheat, generally do not trade with the U.S. or Europe due to increased shipping costs [38]. As such, within the Black Sea market area, Russia and Ukraine command a highly privileged position. However, this local market is relatively small compared to global trade, which explains Russia and Ukraine's low global trade rank.

Moving away from looking at specific exporters, we want to know how model prices compare to other network measures. Shown in figure 5.12, we calculate the model price as a cumulative distribution function (CDF) using the wheat export market data. This shows the fraction of the exporters that command prices greater than or equal to price p. We can see that on *lin-log* axis the cumulative price distribution decays in a fairly linear fashion, perhaps indicating an exponential distribution. Interestingly, the CDF of the degree distribution (see Chapter IV for details of how to draw degree distributions) follows a similar exponential trend. Note that this is only using data from a single year (2015), so it is not conclusive enough to draw any connection between the distributions. However, these results do fit with observations made by Kakade et al. [86], who found that the price distribution of several simulated networks closely resembled the degree distributions of the networks from which they were derived.

We might expect, therefore, that the export price commanded by an exporter would correlate to that specific exporter's degree. We look at this relationship more closely in figure 5.13 (*left*), where we plot each exporter's local equilibrium price p and a function of exporter degree k. We see that whilst there is generally a monotonic relationship for high values of k, this relationship breaks down at smaller degree values where there can be significant price variation across similar degree exporters. It is likely that there is some agreement between model prices and exporter degree for high values of k, because there is a limit on global competition. Recall that

www.manaraa.com

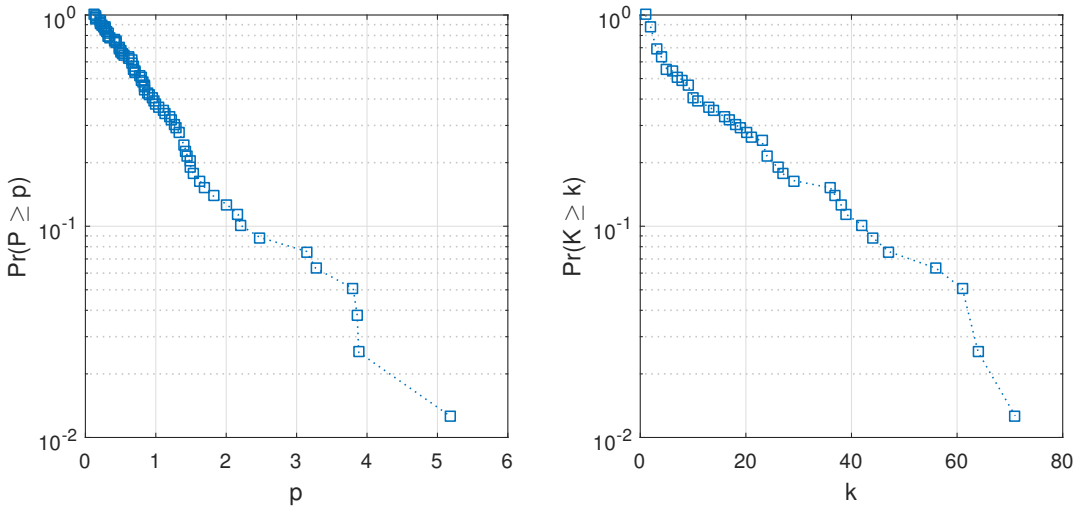


Fig. 5.12 (*Left*) Cumulative distribution of model prices (network data from UN Comtrade database [141]). (*Right*) Cumulative distribution of network degrees. Both distributions show some similarity on *log-lin* scale.

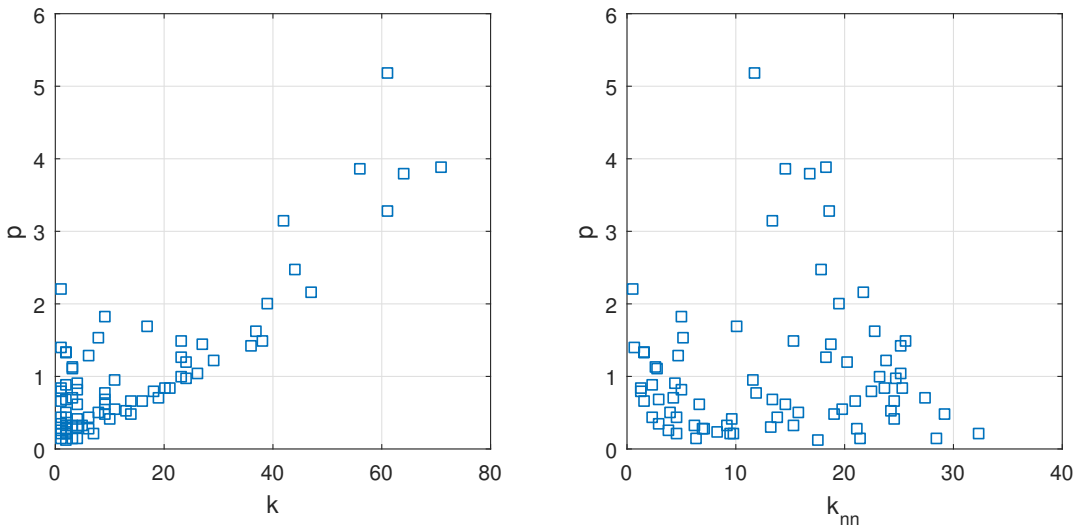


Fig. 5.13 (*Left*) Model prices as a function of country degree. (*Right*) Model prices as a function of country average nearest neighbour degree (network data from UN Comtrade database [141]).

the local equilibrium price is also a function of an exporter's nearest neighbour's degrees (i.e., their closest competition). While there are 79 exporters in the international wheat market, only ~ 30 trade on a truly global scale. The others, as we have seen with Russia and Ukraine are somewhat limited to local areas of the market. This places a firm limit to the level of global competition countries like the U.S., Canada, and E.U. states face across the market. Those countries that can trade with more than this number (i.e., those exporters with $k > 30$), can extract a price premium based on their commanding position. Hence, the increasing correlation

www.manaraa.com

between prices and degree in the upper left section of figure 5.13 (*left*). The plot in figure 5.13 (*right*) shows the *average* nearest neighbour degree for exporters that command prices p. We see no discernible relationship between these two variables. However, we can see that the highest average nearest neighbour degree is $k_{nn} = 32$, which again highlights the limited competition for very high degree exporters.

We have seen that exporters can expect different prices across the network. However, we want to measure this global variation in prices more formally. If the price distribution (see figure 5.12) were normally distributed we could just use the standard deviation, but with the non-normal distribution of prices, the so-called 'Gini coefficient' measure is more suitable. The Gini coefficient was popularised as a measure of income inequality in the fields of economics and international development. However, inequality can be calculated for any distribution—not just for income. For example the global inequality between nation states has been estimated using the Gini coefficient on measures of GDP (Gross Domestic Product). The Gini coefficent gives a measurement between 0 and 1, with 0 representing perfect *equality* and 1 representing perfect *inequality* (e.g., all GDP is concentrated in a single country, while all others have none). However, in practice, these extreme values will never be reached [31]. In the context of a networked market, we interpret the Gini coefficient as a measure of price inequality, which is equivalent to the price difference between all pairs of exporters in the market. We could also interpret the Gini coefficient as a measure of wealth inequality, where wealth would refer to the amount of cash each exporter would make after the market clears at equilibrium prices. Recall however, that an exporter is only given 1 unit g_j of wheat to sell, so the end wealth of each exporter, after equilibrium is reached, is equal to the local price p_j (i.e., wealth $w_j = g_j \cdot p_j$, where $g_j = 1 \quad \forall j$). Therefore, these two interpretations are functionally the same. More specifically, if p_i is the local equilibrium price of exporter i, and there are n exporters in the market, then the Gini coefficient for price inequality is given by,

$$\text{gini coef.} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |p_i - p_j|}{2 \sum_{i=1}^{n} \sum_{j=1}^{n} p_j}. \tag{5.3}$$

We can see the Gini coefficient is equivalent to the *mean absolute difference* between all pairs of exporter prices in the market, divided by half, and then normalised to a scale between $[0, 1]$.

Using our model, we calculate the Gini coefficient for the international wheat market as 0.41, which is relatively low (i.e., prices are more equal). For reference the Gini coefficient for

www.manaraa.com

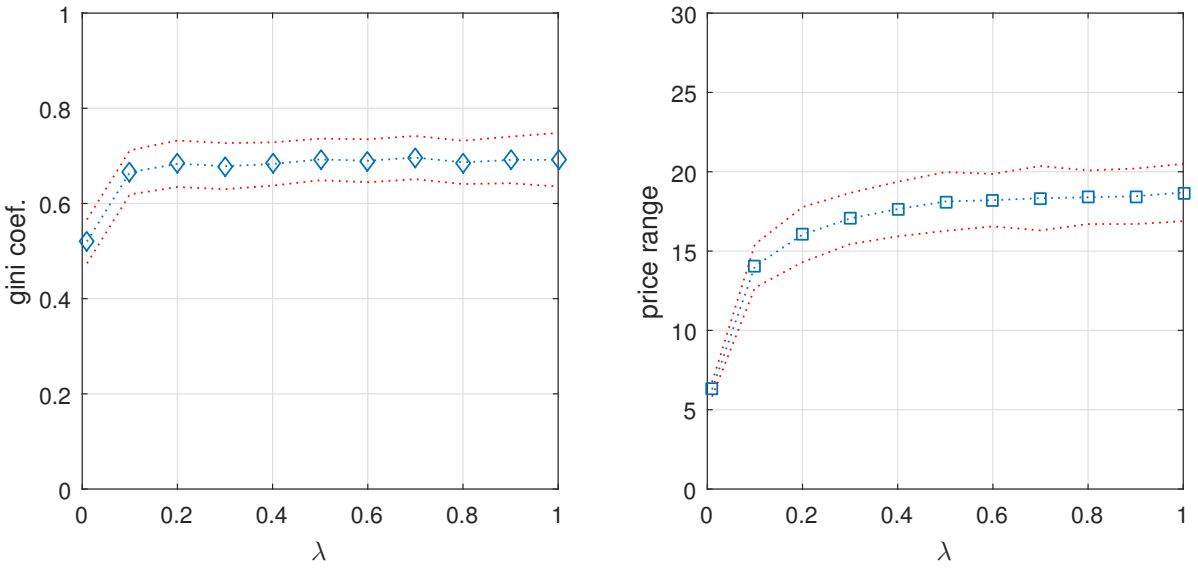


Fig. 5.14 (*Left*) The average Gini coefficient of synthetic markets generated using the constrained growth algorithm with varying values for the growth parameter λ. (*Right*) The average range in model prices as a function the growth parameter λ.

the world GDP has been estimated to be between 0.61 and 0.68 [31]. By itself this value is not going to tell us much about how the network structure of a market affects the inequality between exporters. So to shed light on this issue, we simulate several synthetic markets using the network growth algorithm developed in Chapter IV. Recall that the model parameter of this algorithm, λ, denotes the probability of adding a new vertex to the network each time period, with time here referring to the number of edge-attachment events. We apply the graphical exchange model over the network structures generated by the algorithm to derive local equilibrium price distributions for each market. These distributions are then used to calculate the Gini coefficient for each market. In figure 5.14 (*left*) we plot the Gini coefficient from each market as a function of the model parameter λ, that was used to generate that market. In figure 5.14 (*right*) we also plot the range of equilibrium prices (i.e., $p_{\max} - p_{\min}$) as a function of the parameter λ, so we can see the effect in terms of real price changes. We can see that across both plots the variation in prices dramatically increases between $0 < \lambda < 0.2$. After which, both price inequality and range seem to level off, with only a small increase between $0.2 < \lambda < 1$ in the range of model prices.

So the question becomes, what structural change is happening to the underlying network that may be responsible for this increase in price variation? Recall that the main feature of the network growth algorithm was that lower values of λ (note that here, λ acts as a constraint on extrinsic network growth), has the effect of homogenising vertex degrees across the network.

www.manaraa.com

We find that setting $\lambda \leq 0.1$ produces networks broadly equivalent to a random network, in that edge attachment rates were uniform and the resulting degree distribution was exponential. Alternatively, setting $\lambda = 1$, resulted in networks that were similar to those produced using the Barabási-Albert (B-A) model [15], such that the network developed into a scale-free topology with a power-law like degree distribution. Looking again at figure 5.14, we can say that the structural change across the horizontal that might explain the Gini coefficients, is that the degree distribution of the networks are becoming more heterogeneous as λ tends to 1. One can imagine that this heterogeneity between exporter connectivity could naturally lead to situations where a few exporters, at the high end of the degree distribution, might find themselves in a privileged position, not only being more connected to the global market, but also facing less competition from other high degree exporters. It should be noted that we find most real-world export markets between $0 < \lambda < 0.06$, which would indicate relatively similar market prices. However, we can see that the Gini coefficient is very sensitive to changes in λ precisely along the parameter values where these markets lie. Thus, any change in growth rates due to increasing globalisation, for example, could dramatically increase the inequality between exporters in these markets.

5.6 Conclusion

In this study, we have built on previous research to show how a country's export performance can be influenced by its position in the market's network structure. We analysed how network asymmetries between countries in global markets can lead to varying prices for the same good. We find that increasing an exporters degree (i.e., the number of importers that a country can trade with) also increases the prospects of obtaining better prices. However, we also find that increasing competition in the market, measured by the average nearest neighbours degree, rapidly reduces the maximum price an exporter can command. When measuring prices across networks, we find that price variation increases as network degree distributions become more heterogeneous. Testing these theories on real-world trade data reveals that the network structure of commodity markets places them in a very sensitive position, whereby small changes to the network's degree distribution could have large implications for the equality of prices across the market.

www.manaraa.com

Chapter 6

The Risk of Export Restrictions to Globally Traded Food-Commodities

6.1 Introduction

This chapter contributes to a small, but growing literature on how the network structure of international trade is becoming a crucial element in understanding global food security. From the work reviewed in Chapter III, we have seen how the global trade in commodities has created a complex network of interdependent supply chains. This network is especially important in maintaining the supply of agricultural and food-related commodities. However, as more countries have become reliant on this network, the more vulnerable it has become to systemic risks and shocks [48]. Research has shown that *export restrictions* play a key role in exacerbating these risks, evidenced by the recent episode of extreme food-price volatility, commonly referred to as "the global food crisis". From a network perspective, a country's susceptibility to these risks is directly related to its level of integration into the world trade system's structure [87]. Here, we develop a network theory approach to analyse the effect of export restrictions in three major food-commodity markets. The commodities—*wheat*, *rice*, and *maize*—are used as examples throughout the chapter, as they collectively provide the world's population with most of its calorific energy [151]. The network method allows for the extraction of relevant structural information about the effect of export restrictions on the network as a whole [122], whilst also identifying the relative importance of specific countries [153]. We use this analysis to develop a composite risk index that assesses the risk posed by export restrictions to food-importing countries based on their position in the network. We interpret these risks in terms of developing a policy of strategic reserves and provide a case study example showing how the network metrics can be applied to mitigate a country's food security risk. The chapter

www.manaraa.com

is structured as follows: in sections 6.2 and 6.3, we give a brief overview of the global food crisis and present the evidence regarding its origin. We then outline the role export restrictions played in the crisis in Section 6.4. In Section 6.5, we detail the model and methods used in this study. The results of our analysis are then presented in Section 6.6, and in Section 6.7, we develop a case study example to interpret these results when applied to food security.

6.2 Background to the Global Food Crisis

After decades of relative stability, several major food-commodity markets have experienced persistently volatile prices over the period 2007-2013. The rapid increase in the price of basic food-commodities and the often violent protests that came in their wake have led academics to call these events a "global food crisis" [134, 151]. According to FAO (Food and Agriculture Organization of the United Nations) data [49], world food prices began a persistent upward drift starting in 2002. By the end of 2006, the food price index shown in figure 6.1 had moved almost 25% higher than the average price from the previous decade (i.e., 1990–2000). Price rises then started to accelerate, jumping to 60% by the end of 2007 [150]. Six months later, in June 2008, the price index peaked at around 76% higher than average. Price changes in the cereal-only index were even more dramatic, rising 208% above the index average by June 2008 (see figure 6.2). Food commodity prices then plummeted just as fast as they had risen. The food price index started falling after only two months, ending at 22% of its peak by the end of 2008. Cereal prices also fell precipitously, ending at 32% of peak prices. The rapid rise and fall in prices is characteristics of a classic price bubble, with the overshooting of prices leading to an inevitable violent correction [74]. Since 2008, prices have exhibited an unusual and sustained volatility, with further price spikes in 2010 and 2013 (see figures 6.1 and 6.2).

The degree to which the food price index jumped in 2008 is all the more striking given that aggregated price indices tend to smooth out the ups and downs of individual commodity price changes [151]. In specific markets, prices did not always change synchronously. Wheat and maize prices rose first, in June and September 2007 respectively. Rice prices did not spike until later, at the beginning of 2008, but subsequently rose much faster and further. Wheat prices were also the first to peak in March 2008, at 204% higher than the index's average. Rice prices peaked at highs of 378%, in May 2008. While maize prices peaked at around 172%, in June 2008. Despite the rapid falls, wheat prices were still 38% above the index average at the end of 2008. Maize prices were up 50%, and rice prices were still up 160%. The speed of those price

www.manaraa.com

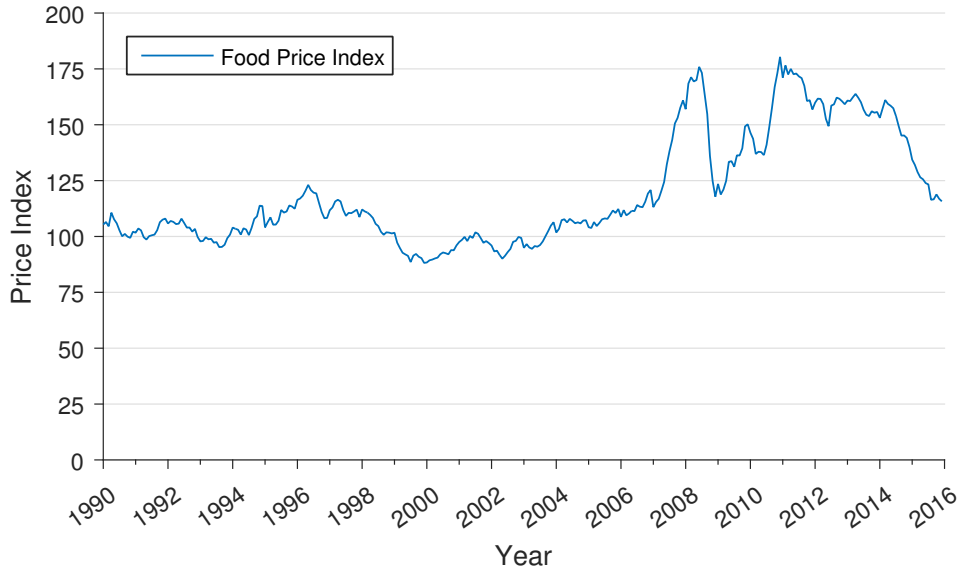


Fig. 6.1 UN-FAO monthly food price index (1990–2016). The food price index consists of the average of five food-commodity price indices (Meat, Dairy, Cereals, Oils, Sugar), weighted with the average export shares of each of the groups in 2002–2004. The index is then deflated using the World Bank Manufactures Unit Value Index (MUV), and normalised such that 2002–2004 average = 100. Data source: [49].

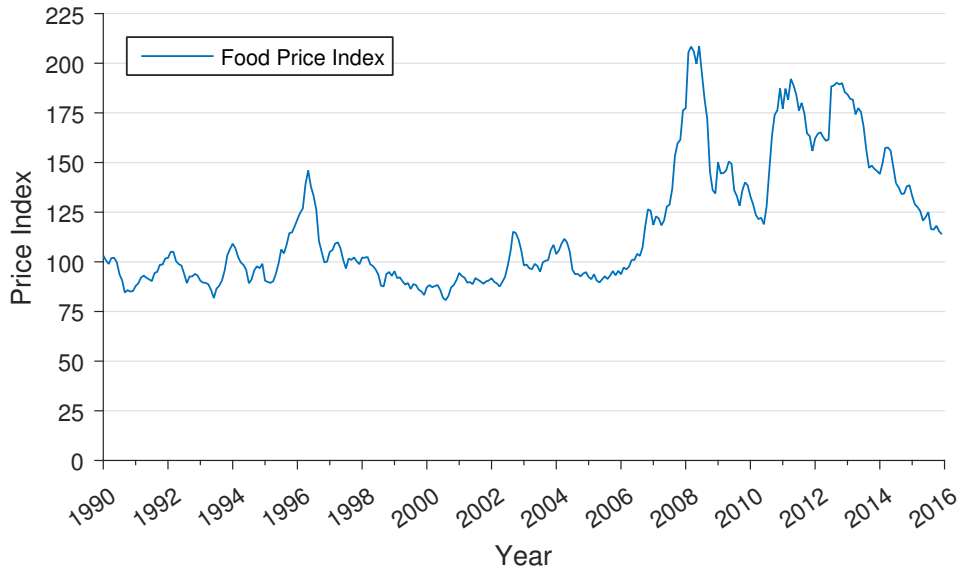


Fig. 6.2 UN-FAO monthly cereal price index (1990–2016). The cereal price index is compiled using an average of 10 different wheat price quotations, 1 maize export quotation and 16 rice quotations. The quotations are combined by weighting each commodity with its average export trade share for 2002–2004. The index is then deflated using the World Bank Manufactures Unit Value Index (MUV), and normalised such that 2002–2004 average = 100. Data source: [49].

www.manaraa.com

changes was also an important factor in the crisis [74].

The speed of price changes in a given commodity is often referred to as market volatility. However, there is no single way of measuring volatility, as there is an element of subjectivity in specifying a price volatility model. The simplest way is to base a measure of volatility on the standard deviation of an observed random variable, in this case, commodity price changes between years [112]. Volatility measured in this manner is referred to as annualised historical volatility [112]. The main problem here, is that most time-series datasets follow a trend (e.g., due to rising inflation). If not removed, these trends will be accounted for in the volatility measure [112]. However, de-trending requires a judgemental trade-off between the attribution of variability to the trend itself (for example, the rate of inflation) and to variation about the trend (i.e., the true volatility in prices) [65]. We calculated the values shown in figure 6.3 without de-trending, since the period covered is relatively short, but a reasonable model would perhaps be 3% year-on-year growth in food and cereal prices indices from the year 2000 onwards. Figure 6.3 (**a**) shows the market volatility in the UN FAO food price index, and (**b**) in the cereals price index, as such, they are re-interpretations of figures 6.1 and 6.2. We can see that, in terms of volatility, the food price index had a background level of 2-5% variation in prices each year until 2007, when the level of market volatility increased 3-4 times, topping at 20% in 2008. The cereal index had been far more volatile prior to the crisis than food prices in general, averaging 5-15% between 1990 and 2006. It also saw higher volatility during the crisis, with 25% and 30% changes in 2007 and 2008. While, annualised historical volatility generally follows the price index they track, high volatility in cereal prices in 2010 highlights where measures of volatility and prices changes can diverge. Here, the price index changed only 39 points across 2010; however, the absolute change in price (lowest price to highest price in a given year) in 2010 was almost double that, at 69 points. This translates to an annual volatility of around 25%.

The impacts of the crisis were especially disruptive for poor consumers, who spend a large share of their income on basic food-commodities [1]. Such large price rises can set in motion a downward spiral of vulnerability, accentuate poverty, and possibly lead to malnutrition [134]. The World Bank (cited in Mundial [103]) estimated that the crisis forced an extra 105 million people into extreme poverty[1], adding to the 883 million people already living below this threshold. Likewise, the FAO [47] food security assessment estimated that the crisis increased the number of undernourished people in the developing world from approximately 858 million

[1] Extreme poverty is defined as living on less than $1 per day [103]

www.manaraa.com

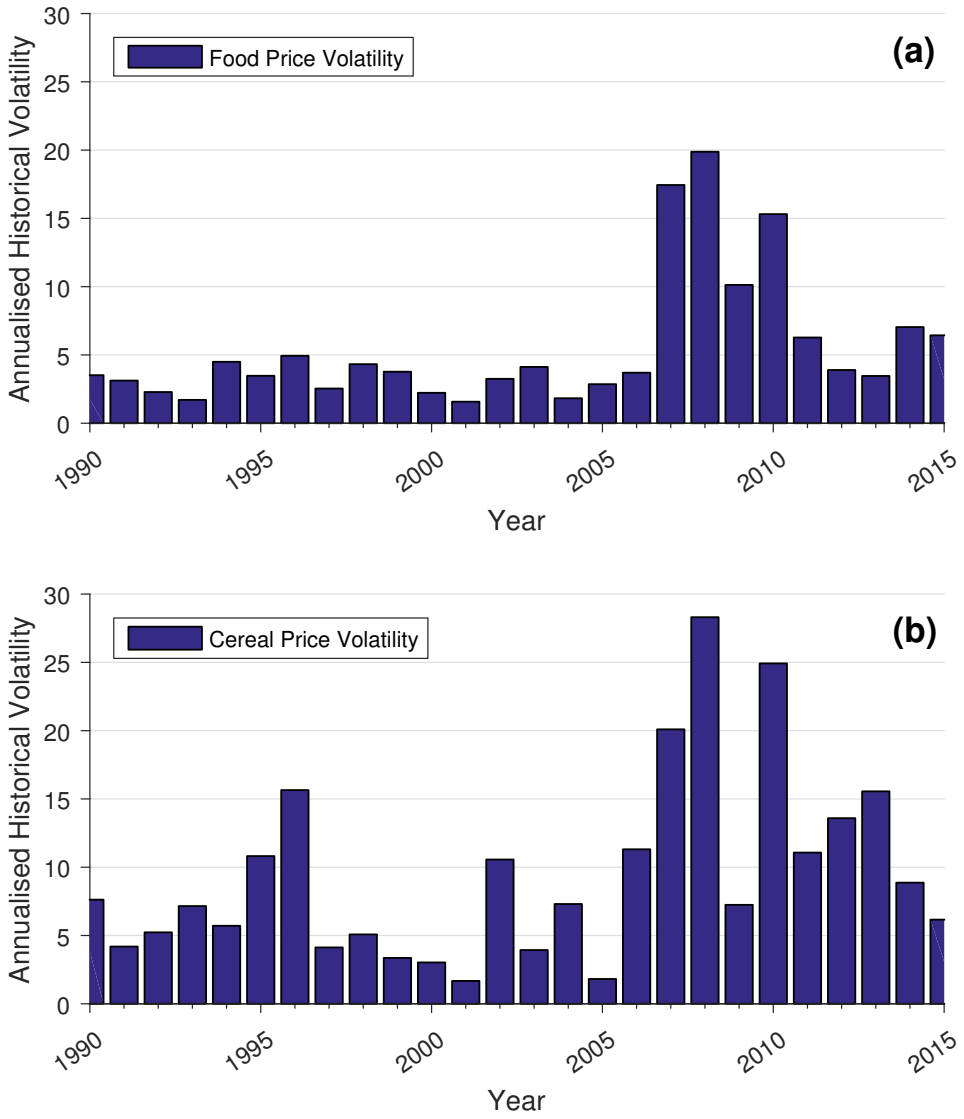


Fig. 6.3 (**a**) Annualised historical volatility in the UN-FAO food price index (1990–2016), and (**b**) volatility in the UN FAO cereals price index (1990–2016). Data source: [49].

www.manaraa.com

to 923 million, totalling an increase of 65 million people becoming food insecure since prices began to rise. The USDA [142] estimate those figures at 849 million before the crisis, and 982 million after[1]. The crisis also reversed progress made by many countries towards meeting the Millennium Development Goals to reduce hunger and malnutrition [140].

6.3 Origins of the Crisis

In the absence of a clear causal event, numerous arguments concerning the origins of the crisis have been proposed. It is often noted that the causes of high commodity prices are the result of a complex set of interrelated factors involving both the long-term fundamentals and short-term shocks [144]. In terms of long-term trends, global population growth and strong economic growth in Asia, undoubtedly contributed rising prices [134]. Both of these have led to rising living standards and changing diets (particularly in increased meat consumption) across the developing world, which, in turn, has led to an increase in demand and trade in agricultural commodities globally [134]. While the general shift towards higher prices can partly be explained by these long-term trends, it is also likely that changes in several other key factors just prior to 2008, made the crisis all the more severe [74]. Ordered by the weight of evidence in the literature detailed below, the most plausible of these factors are as follows:

- Historically low food-commodity stocks.
- Supply shocks caused by droughts and other extreme weather events.
- The diversion of grains to biofuel use.
- The petroleum price pass-through effect.
- The depreciation in the value of the U.S. Dollar

We review and critique the evidence for each factor below.

6.3.1 Low Stocks of Food-Commodities

Perhaps the least controversial causal factor proposed, is the decline in global stocks of many staple food-commodities. Scholars are almost unanimous in pointing to the expectation of low

[1] It is worth noting that the indicators outlined are based on relatively simple predictive models using data collected prior to the crisis [1]. Little work has been done to validate these estimates after the event; therefore, they should be interpreted with some caution.

www.manaraa.com

stocks in both 2007 and 2008 as having either a direct effect on rising prices, or as being a prelude to the situation [1, 138, 101, 150, 111, 75, 134]. Stocks of grains, such as wheat, rice, and maize can be taken as a direct indicator of food security in both self-sufficient and net-food importing countries. These three commodity markets typically have inelastic demand, as the majority of consumers depend upon grains as their staple food source. As such, consumers will continue to purchase at high prices, even if that means giving up other expenditures, such as health and education [150]. Supply of these commodities can also be somewhat inelastic as there is a seasonal time lag between farmers receiving new price signals, and their ability to change levels of production. However, when stocks are added to the production supply, the total supply in the market becomes more elastic, as any increase in consumption (or decrease in production) can be compensated for by the release of stocks [1]. Stocks can also limit the effect of speculation on market prices, as they can be unloaded to negate the position speculators may take in betting on rising prices [111, 150, 75].

Theoretically, when stock levels are low, relatively small changes in supply can result in rapid price changes [111, 150, 75]. If stocks run out completely, then market demand becomes identical to the consumption demand. At this point, any further shocks will have an extreme effect on prices [1]. According to standard economic theory, we might expect that when stocks are high, prices should be low and stable. Conversely, when stocks levels decline, prices will rise and be volatile. We can assess these assumptions by comparing a measure of global stocks, called the stock-to-utilisation (STU) ratio against the global commodity price index. The stock-to-utilisation (STU) ratio indicates the level of carryover stock for any given commodity as a percentage of the total demand or use. It is often used by traders as a rough measure of the prevailing supply-demand conditions [137].

More formally, the STU ratio can be described as follows,

$$\text{STU} = \frac{\text{Beginning Stock} + \text{Total Production} - \text{Total Use}}{\text{Total Use}} \times 100. \tag{6.1}$$

and can be simplified by consolidating the numerator to,

$$\text{STU} = \frac{\text{Carryover}}{\text{Total Use}} \times 100. \tag{6.2}$$

From equation (6.1), beginning stocks simply represent the previous year's ending stock of a particular commodity, either on a per country basis, or the total stock available globally. Similarly, total production represents the total grain produced in a given year and total usage is

www.manaraa.com

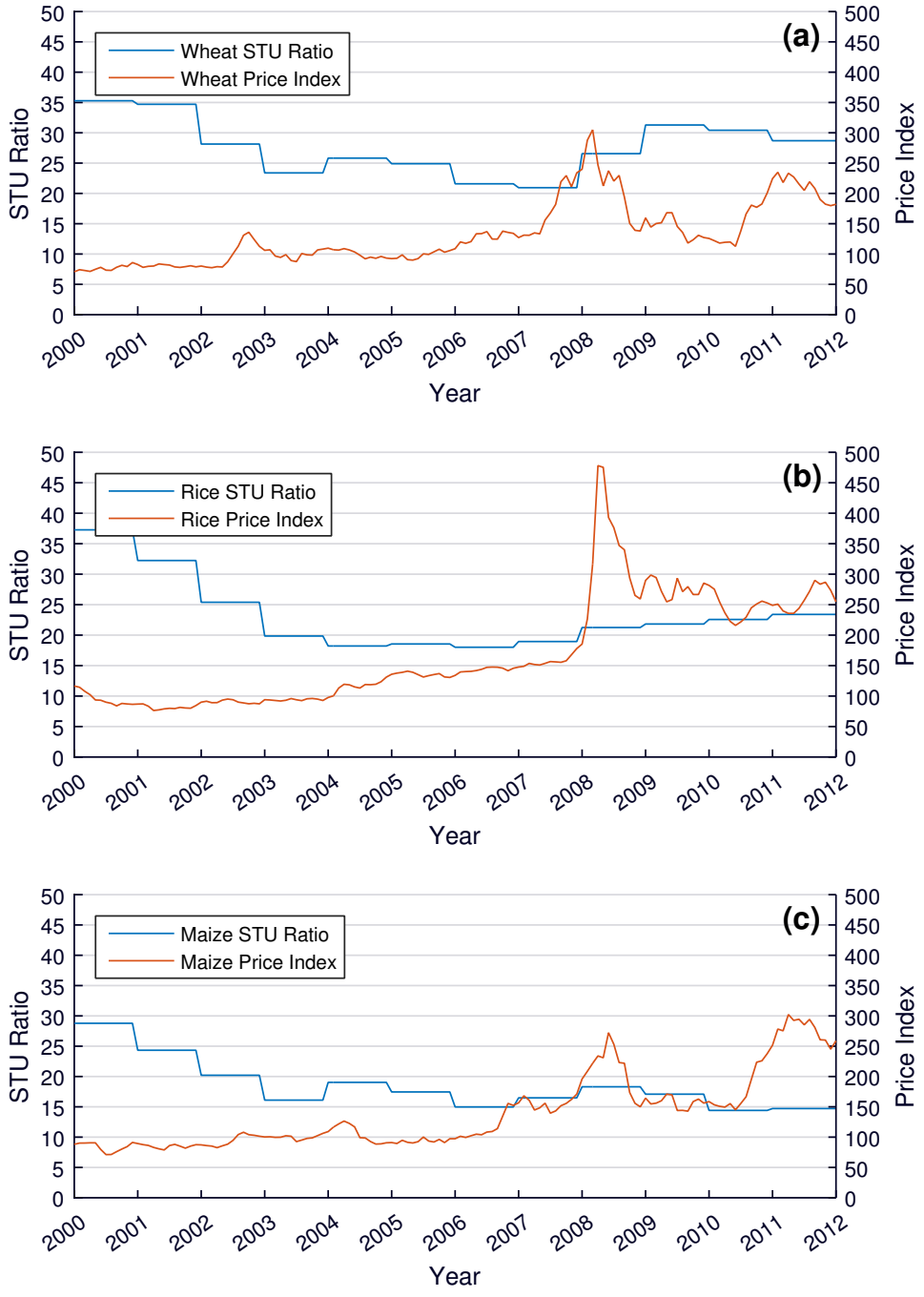


Fig. 6.4 On the left axis, global stock-to-utilisation (STU) ratios in (**a**) wheat, (**b**) rice, and (**c**) maize, and on the right axis, the global price index of each commodity (years 2000–2012). Data source: [143].

www.manaraa.com

the sum of all the end uses in which the stock used (i.e., consumption, waste, and exports). From equation (6.2) the carryover is the amount of a commodity leftover and stored for consumption the following year. The carryover stock divided by the total usage can be expressed as a ratio, i.e., the stock-to-utilisation ratio. This ratio can be compared between years to indicate whether current stock levels are low or high relative to past performance. Some trades use the STU ratio to estimate both the direction of a price trend, as well as the size of those changes [137].

Historically, food-commodity markets have been characterised by a high-stock/stable-price regime, infrequently punctuated by low stocks and volatile prices. From this historical record, some benchmark ratios have been established for various commodities. For example, STU ratios for wheat under 20% has typically led to strong price rises. In the case of maize, price rises tend to require much lower STU ratios of around 12% [137]. These benchmarks are evidenced by experiences in the 1974 food crisis. Prior to 1974, wheat stocks declined from a peak of 41% STU ratio in the 1960s to only 21% by 1972. At this stock level, a modest reduction in world wheat production is thought to have caused the annual price index to more than double [74]. Similar events played out in the global maize market, where stocks declined from highs of 31% STU ratio in 1960 to just 12% before the 1974 crisis [116]. However, unlike wheat and maize, rice stocks rose continuously from the 1960s through to the 1980s due to the 'Green Revolution', which saw agricultural production levels rise significantly due to the introduction of high-yield grains, hybridised seeds, along with new fertilisers and pesticides [73]. As such, rice stocks were not seen as a factor explaining the 1974 crisis [74].

As we can see in figure 6.4 (**a**), (**b**), and (**c**), stocks of all three commodities declined preceding the recent 2008 crisis. From highs of around 30-35% STU ratio in 2000, wheat and maize fell to 20% and rice fell to 15%. According to FAO [46] data published in July 2008, utilisation exceeded production in all years since 1999, bar 2004 when the harvests were especially good. Headey and Fan [74] suggest that poor weather conditions and bad harvests in 2000, 2002 and 2003 are the cause, leading to the depletion of stocks globally. Although the better harvest in 2004 was enough to replenish stocks that year, a severe drought in Australia and other weather events led to another poor harvest in 2006, around the time the price rises began (see Section 6.3.2 for further discussion). Stock levels were also thought to have declined prior to the crisis due to policy decisions outside the scope of weather related stock adjustments. China in particular, is thought to have reduced grain stocks in the 1990s and early 2000s, as the benefits were no longer deemed to outweigh the wastage costs. In a similar vein, the advent of just-in-time inventory management was thought to have made the use of large stockpiles

www.manaraa.com

obsolete [74].

Statistical analysis by Piesse and Thirtle [111] shows that the aggregate price index of cereals (seen in figure 6.2) moves in the opposite direction to global cereal stock-to-utilisation levels, implying that changes in cereal prices can be explained by changes in stock levels [111]. In figure 6.5, we expand on this analysis by disaggregating the cereal price index into it constituent components (wheat, rice, and maize) and estimate the relationship between each commodity's STU ratio and the respective commodity price index over the period 1992 through 2015. As expected, there is a general negative relationship between the STU ratio and the respective commodity price index. That is, low STU ratios are associated with higher prices and vice versa. This is evidenced by the negative correlation coefficients (r values) across all three commodities (see figure 6.5 **d, e, f**). We can see that in each commodity market, a regression model with a lagged explanatory variable (i.e., where price is measured at time t and the STU ratio is measured at time $t-1$), (seen in the top row of figure 6.5 **a, b, c**) results in the most robust model in terms of both goodness-of-fit, measured by the R^2 value, and the p-value measuring statistical significance, i.e., that the slope of the regression model is significantly different from 0. Conversely, where the regression models are forward lagged such that the STU ratio is measured at time $t+1$ (seen in the bottom row of figure 6.5 **g, h, i**), the relationship is found not to be statistically significant in any market. This is a good indication that price changes follow changes in STU ratios, and not the other way round. Notwithstanding these results, this relationship is far from established, as even the best models can only explain between 33-49% of the commodity price changes. For example, looking at the maize model in figure 6.5 (**c**), it might be more appropriate to use a non-linear regression model, similar to that used by Good and Irwin [70], who achieved a better fit when estimating the relationship between U.S. soybean STU ratios and prices using a model in the form,

$$y = a(1/x) + b, \tag{6.3}$$

such that, the regression line becomes steeper and steeper as the STU ratio declines, and *vice versa*.

Despite the relative weakness of the regression modelling, these data do show that, generally speaking, the theory behind STU ratios and commodity prices is correct. We see across all commodities and models that when stock levels are high relative to consumption, prices tend to be both low and non-volatile, whereas at low STU ratios, prices are found to be both high

www.manaraa.com

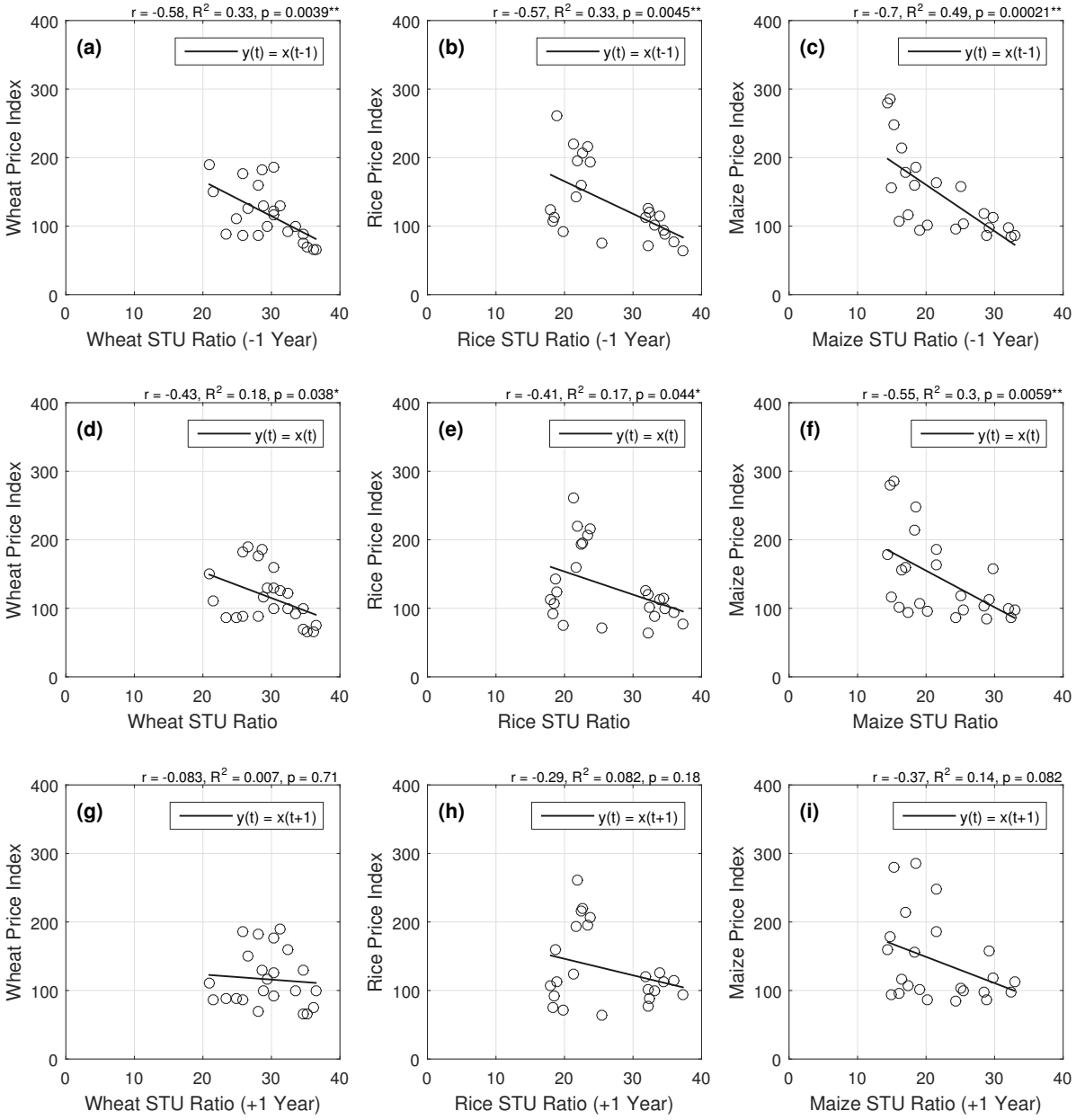


Fig. 6.5 Correlation of wheat, rice, and maize stock-to-utilisation (STU) ratios and price indices (years 1992–2015). Sub-plots (**a, b, c**) show the commodity price index against the previous year's STU ratios, with (**d, e, f**) against the same year's STU ratios, and (**g, h, i**) against the following year's STU ratios. Statistical significance is indicated by (*) where $p < 0.05$ and (**) where $p < 0.01$. Data source: [143].

www.manaraa.com

and low, suggesting higher levels of volatility. Hepburn [77] argues that these results indicate that low stocks are a necessary but not a sufficient condition of price rises. Rather, there needs to be a large adverse supply shock in addition to low stocks to trigger crisis conditions. The following two sections examine the supply situation prior to the 2008 crisis.

6.3.2 Droughts and Supply Shocks

As mentioned above, drought and other weather related events have been identified as another major factor underpinning the 2008 price rises and/or stock declines [74]. The period preceding the 2008 crisis was marked by several serious droughts that caused harvest failures in several major grain exporters [109]. Australia was hit the worst, suffering back-to-back droughts in 2006 and 2007 which reduced total grain exports by an average of 9.2 million metric tonnes compared to 2005 [111]. Ukraine was also hit by poor harvests in 2007, which reduced their grain exports by an additional 10 million tonnes compared to the previous year [111].

Mitchell [101] suggests that singling out those countries that were affected by weather related shocks is being somewhat selective. In any given year, some countries are bound to have better than average harvests and others worse. For example, in 2007 the decline in exports from Australia and Ukraine were more than offset by good harvests in Argentina, Kazakhstan, Russia and the US (see figure 6.6). Combined, these countries increased their total grain exports by approximately 22 million tonnes compared to 2006. The net effect was a 4.7% increase in global gain exports in 2007 [111].

To look into this issue in a more systematic way, we disaggregate exports in the global wheat market over the 2000–2011 period, shown in figure 6.6. We can see that during the period prior to the crisis, only in 2003 was the net position of the market significantly down in export volumes, the major factors being drought in Australia and poor harvests in Ukraine. 2007 was also a poor year in exports, primarily due to the continuing drought in Australia. However, strong export volumes in the U.S. made up for losses elsewhere, leaving the net market position down only 1.3 million metric tonnes (small compared to the 130 million tonnes of wheat traded in 2007).

After the crisis, high prices brought about a global supply response, with harvests up 4.9% between 2007 and the end of 2008. Net exports also increased by 7.5 million tonnes in 2008,

www.manaraa.com

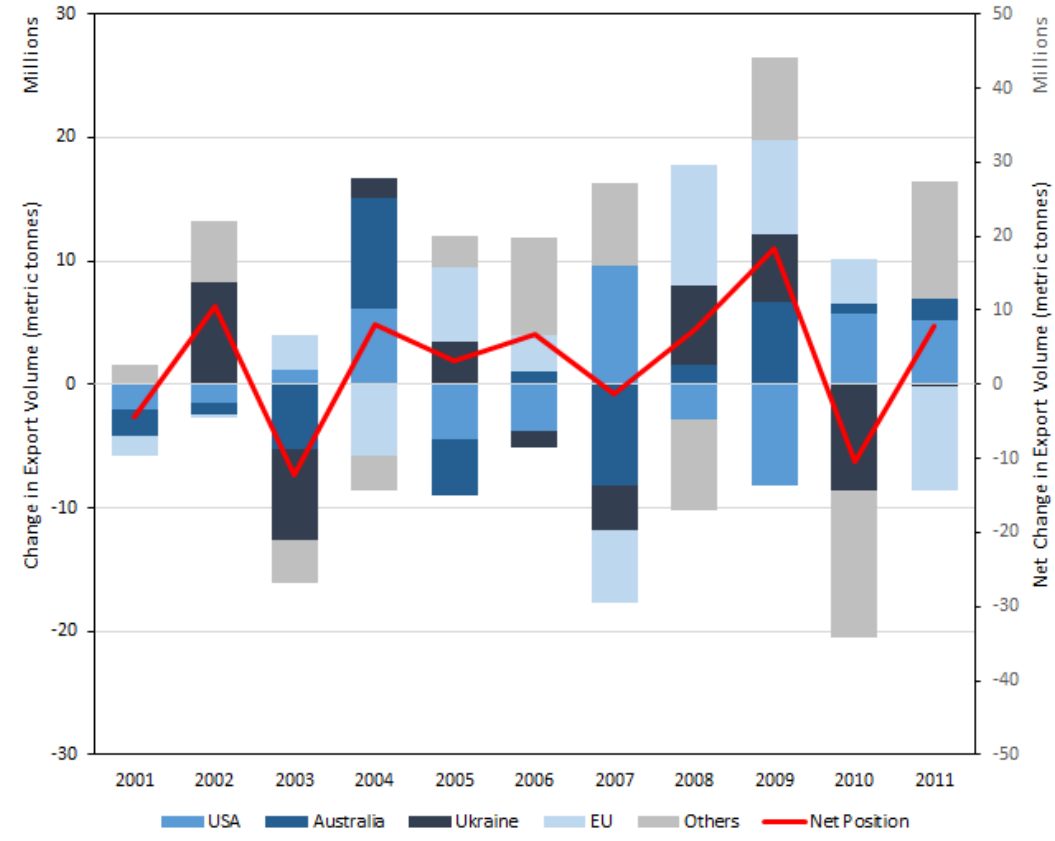


Fig. 6.6 Change in export volume in global wheat markets, 2001–2011 (measured in metric tonnes). Colours show the change in each country's export volume and the red line shows the change in the net position of the market as a whole. Data source: [49].

www.manaraa.com

and 18 million tonnes in 2009 [49]. The post crisis period also saw stock-to-utilisation levels increase by around 10% in wheat markets, 5% in rice, and 3% in maize (see figure 6.4).

6.3.3 Biofuel Production

Biofuel production in the US and EU created a major new source of demand in maize and vegetable oil markets [101, 120, 144, 75]. It is thought that the rapid expansion in production led to significant supply shocks in these markets, as well as knock-on effects in other food-commodities. However, the causality of how biofuel production affects food prices is not entirely straight forward. When oil prices rise high enough (above approx. US$60 a barrel), suitable grains, such as maize, are often diverted from food and feed to biofuel production [119]. With oil prices reaching record highs, the use of maize for biofuels, such as ethanol, grew rapidly from 2004 to 2007, to the point where biofuel production accounted for 70% of the increase in global maize production. In the United States, the diversion of maize to biofuels was approaching 30% of all production, some 81 million metric tonnes by 2007 [143]. Given that the United States accounted for 60% of global maize exports in 2007, changes in U.S. production significantly affected international prices [101]. Mathematical simulations developed by the International Monetary Fund [82] estimates that the diversion of maize accounted for 70% of the increase in maize prices over the period 2006 to 2008 [93]. A similar study by Schmidhuber [119] puts that figure at 60%. At the same time, the area planted for maize also underwent a rapid expansion, increasing by 23% in 2007 alone. This expansion came at the cost of other important food-commodities, leading to a rise in prices across all markets. For example, a 16% decline in soybean planted area in 2007 is thought to have contributed upwards of 40% to the increase in soybean prices in 2008 [93]. In Europe, other oilseeds used in biofuel production displaced wheat for the same reasons [75]. The total impact on global grain prices, estimated using a partial equilibrium model, found a 30% increase across all grains. Individual estimates put maize, wheat, and rice price increases at 47%, 26%, and 25%, respectively [117].

Critics of the bio-fuel theory point out that the diversion of grains to biofuel did not come as a complete surprise in 2007. Rather, the increase in demand was ‘baked-in’ to maize prices well before the crisis. Biofuel mandates in both the US and EU were viewed as quasi-permanent policy commitments, with foreseeable implications for the related growth in demand for grains and oilseeds [150]. As Headey and Fan [75] point out, other biofuel crops such as sugarcane that are not under mandated production, did not experience similar price surges.

www.manaraa.com

6.3.4 Petroleum Price Pass-Through

Several authors have highlighted the impact of rising crude oil prices [75, 1, 101, 151]. Oil is used extensively throughout both the production and transportation of all commodities. Therefore, rising oil prices during the 2007–2008 period arguably provide a strong explanation of commodity-price escalation in both food and non-food commodities [75]. Moreover, agriculture is more reliant on fuel intensive inputs such as fertiliser, which make up a large percentage of farmers' variable costs [1, 3, 151]. According to the International Energy Agency [80] agriculture is second only to transport in the oil intensity of its energy usage, suggesting agricultural input costs could be quite sensitive to pass-through in oil prices. For example, Agricultural production in the United States is almost solely dependent on oil for its energy use. In addition, given that grains are traded in bulk, agricultural prices are strongly influenced by the cost of transportation [75].

Mitchell [101] provides an appraisal of the effect of rising oil prices on two major grain commodities (wheat and maize) using examples from the U.S. First, he finds the percentage of production costs from energy-intensive inputs (e.g., fertiliser, chemicals, diesel, and electricity) comprise 13.4% of the total cost of wheat and 9.4% of maize. Increases in the price of oil, and the pass-through rise in the cost of these energy-intensive factors combined, led to an average increase in the cost of wheat and maize production by 11.5% between 2002 and 2007. Transport costs also increased because of higher oil prices. Mitchell's analysis of farm-to-export prices suggests that transport costs could have added as much as 10.2% to the cost of maize and wheat. Hence, the combined increase in production and transport costs was approximately 22% over the period 2002–2007. Headey and Fan [75] later recalculated Mitchell's [2008] estimates after testing the sensitivity of his assumptions. Their calculations showed a much higher pass-through of rising oil prices, in the region of 30-40% depending on the model used. They also note that the rise in oil prices predates that in food prices, which is a necessary, but not sufficient condition to notion that rising oil prices where a causal factor in food price rises. Baffes [10] conducted a similar study using econometric methods. Using data from 35 internationally traded primary commodities over the 1960–2005 period, their analysis found crude oil price changes had the highest pass-through to fertilisers at 33%, followed by agriculture at 17%. For comparison, metals had a pass-through of only 11%.

In figure 6.7, we present our own analysis on these data by estimating the relationship between the IMF's crude oil price index and the UN FAO food price index over the period 1992 through 2015. We find only a weak positive relationship (where $r = 0.39$, $R^2 = 0.15$) between

www.manaraa.com

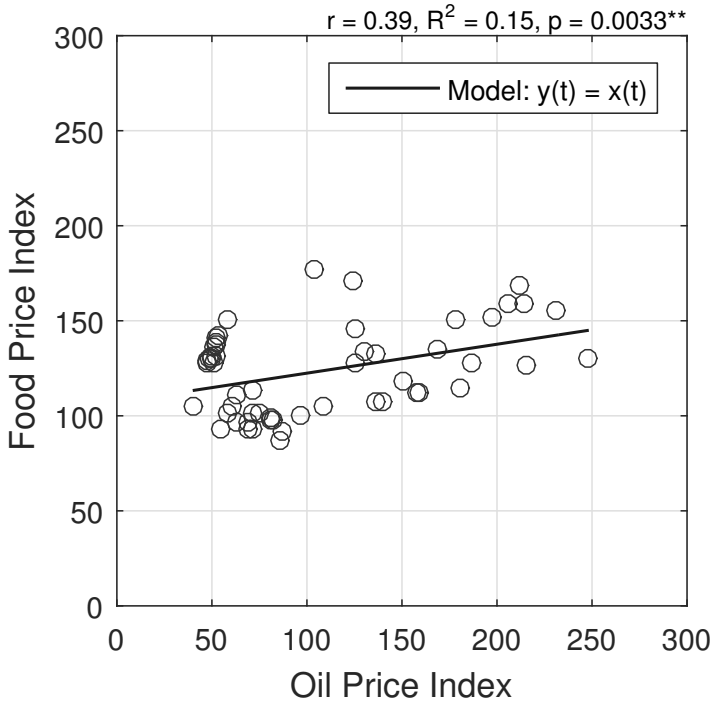


Fig. 6.7 Correlation of the crude oil price index and UN-FAO food price index (years 1960-2015). Statistical significance is indicated by (**) where $p < 0.01$. Data sources: [82], and [49].

the two variables, but one that is statistically significant. Further regression modelling with lagged explanatory variables did not improve the model fit. We also noted, that for specific periods of recent history, the food and oil price indices have been negatively correlated. For example, between 1975 and 1980 the correlation coefficient was, $r = -0.46$, and between 1990–2000, $r = -0.23$. These anomalies cast doubt on the oil price pass-through effect to food-commodity markets.

In addition to the technical analysis outlined above, much has been made of the supposed historical link between oil and food prices [150]. While it is clear that oil prices rose first, and followed much the same trend as food prices during the most recent crisis in 2008 (see figures 6.8 and 6.9), the historical record is far more nuanced than perhaps many authors have suggested (see Headey and Fan [75] for examples). The precedent event, which started this line of reasoning, was the 1974 food crisis. Much like the most recent food crisis, decades of relative stability preceded the shock in 1974. The 1974 food crisis saw prices increase by 75% in just two years (see figure 6.8). 1974 also saw the culmination of another major international crisis, the so-called "first oil shock" caused by OAPEC's (the Organisation of Arab Petroleum Exporting Countries) oil embargo against the U.S. due to its involvement in the

www.manaraa.com

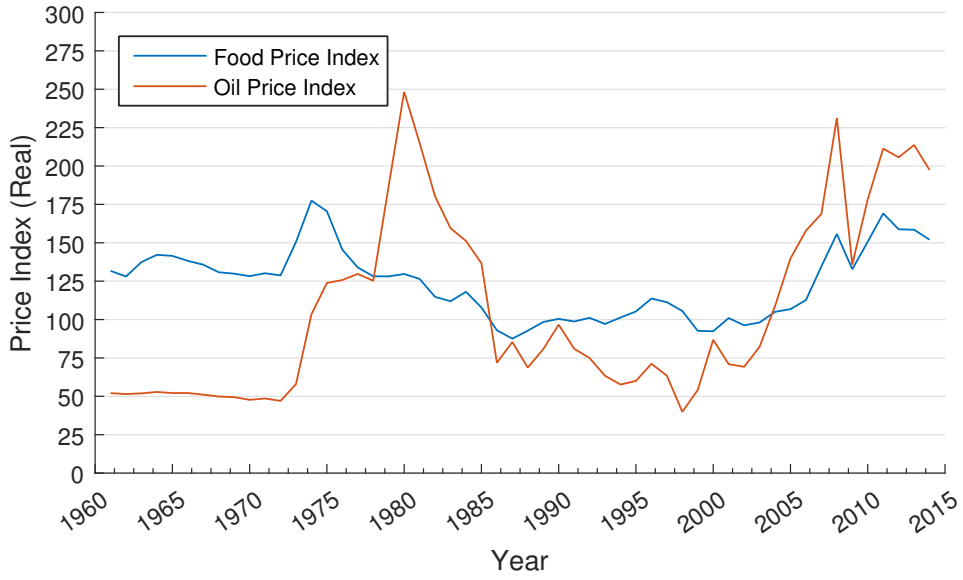


Fig. 6.8 Annual Crude Oil Price and UN FAO Food Price Index from 1960 to 2015. Real prices are adjusted for inflation to March 2015 prices using the Consumer Price Index (CPI-U) as presented by the Bureau of Labour Statistics. Nominal prices are historical free market (stripper) oil prices of Illinois Crude. Data sources: [82], and [49].

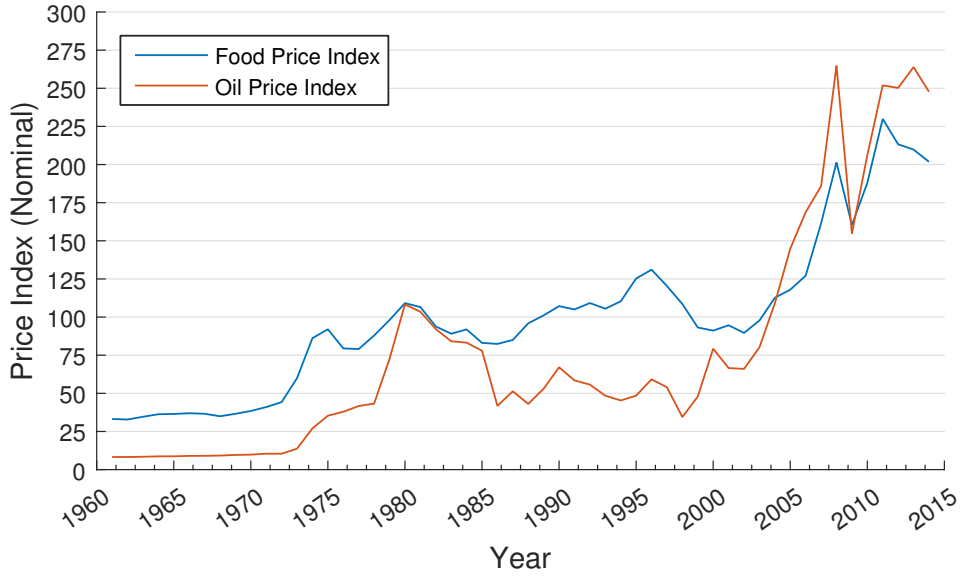


Fig. 6.9 Annual Nominal Oil Price Index and Nominal Food Price Index. Nominal Oil prices are based on historical free market (stripper) oil prices of Illinois Crude. Data sources: [82], and [49].

www.manaraa.com

Yom Kippur War. Prior to this shock, oil prices had also been stable since the end of the Second World War, rising less than 2% per year[1]. Owing to the embargo, world oil prices jumped from around US$3 per barrel to over US$20 (nominal prices). These two events seemed to coincide, and the higher oil prices have long been proffered as a causal factor in the first food crisis.

However, a closer inspection of the data shows that world food prices peaked before oil in 1974. Especially in specific food-commodity markets such as soybeans, which had exports restricted to control domestic food inflation in the U.S. World oil prices on the other hand, only peaked in 1977, before going on to reach new highs in 1980. The 1979 Iranian Revolution caused this second jump in oil prices (referred to as the "second oil shock"), which lead to the suspension of oil exports from Iran (a major exporter). If these increases in the cost of oil were passed-through to food prices, as per theory, we should expect a further food price crisis in 1980. While nominal food prices remained high throughout the 1980s (see figure 6.9), in real terms, food prices went into decline after 1974, and remained low for the best part of two decades. This was due primarily to the effects of the "Green Revolution", which saw large increases in agricultural productivity in the developing world from the introduction of fertilisers and new breeds of crop.

Wright [151] has also identified some problems with oil price pass-through effect in relation to fertiliser costs. He argues that pass-through to fertiliser prices would only affect food production if it caused farmers to cut back on fertiliser use. However, during the 2008 crisis, fertiliser producers like Saskatchewan Potash were operating at capacity and generating large profits [151]. This is because fertiliser has a multiplier effect on farm profits. It is a fairly simple calculation for farmers to make, purchasing fertiliser at 2 or even 3 times the normal rate can still help turn a profit, if the value of the crops grown have also risen. Wright [151] also notes that even if fertiliser inputs had been cut in response to high prices in 2008, their effect could not have been seen until the next harvests, typically almost a year later. However, as figure 6.6 shows, there were record exports in 2009, indicating good harvests that year.

6.3.5 Decline in the value of the US Dollar

Abbott et al. [2] have argued that the depreciation of the U.S. dollar prior to 2008 was also a major factor in food price rises, interpreting the crisis as a predominantly macroeconomic phenomenon, rather than a supply or demand problem. The U.S. dollar began to lose value

[1] Largely due to the Bretton Woods system, which effectively collapsed after 1973

www.manaraa.com

in 2002, first against OECD country currencies, and later against many developing countries' currencies [111]. If the U.S. dollar loses value relative to a currency of an importing country, then that country will face increased costs when importing U.S. goods. Given the U.S. is a major exporter of many agricultural commodities, especially grains such as wheat and maize, then those costs could have risen by up to 63% between 2002–2007 [2].

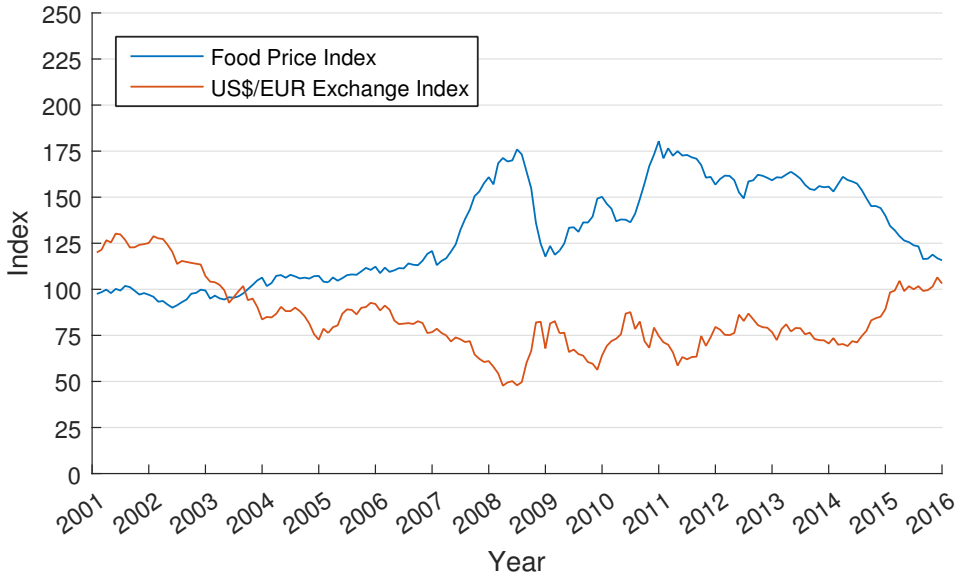


Fig. 6.10 UN-FAO monthly food price index and monthly exchange rate index of U.S. Dollars to Euros (2001–2016). Both indices are normalised such that 2002–2004 average = 100. Data sources: [49], and [133].

However, the effect on food price rises caused by US dollar depreciation remains an area of contention [111]. Abbott et al. [2] are alone in arguing that it was the dominant factor. Gilbert [64] argues that while there is undoubtedly a relationship between the value of the US dollar and food prices, this link is not well understood. He suggests that on this occasion, the depreciation of the dollar was not sufficiently large or general to be a major explanatory factor. For example, while the US dollar fell by 50% over the period 2004–2008 relative to the Euro (see figure 6.10), it changed hardly at all relative to key Asian currencies.

In addition, Mitchell [101] argues that the food price changes prior to the crisis do not closely correlate with the falling U.S. dollar. He states that "from January 2002 to December 2004, food prices rose 24% and the dollar fell 34% against the Euro. Then, from January 2005 to December 2007, food prices rose 65% and the dollar fell 10% against the Euro" [101, p. 11]. Thus, the US dollar fell less when food price increases were reaching their peak [111]. From

www.manaraa.com

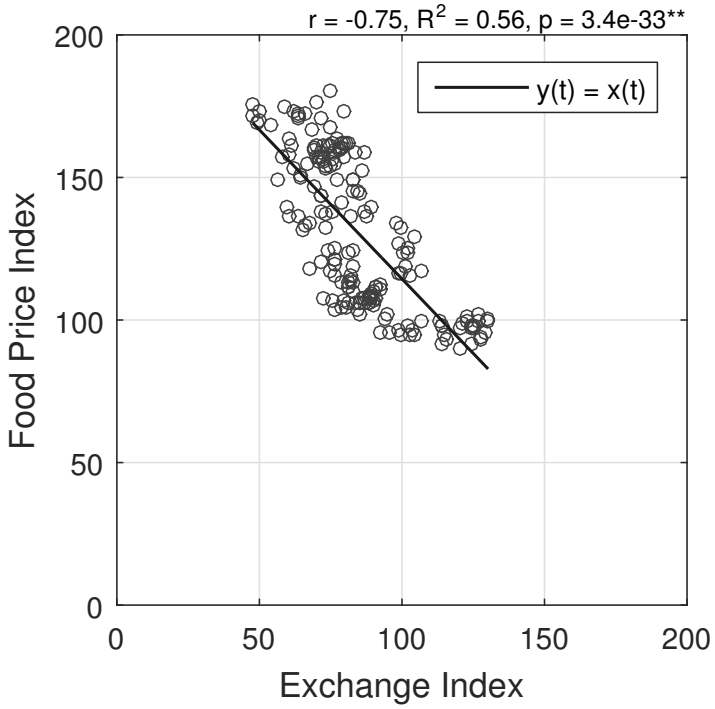


Fig. 6.11 Correlation between UN-FAO food price index and the exchange rate index of U.S. Dollars to Euros (2001–2016). Statistical significance is indicated by (**) where $p < 0.01$. Data sources: [49], and [133].

these findings, Mitchell [101] estimated that only 20% of the rise in food prices could be due to U.S. dollar depreciation.

In figure 6.11, we estimate the relationship between the monthly U.S. Dollar/Euro exchange rate index and the UN FAO food price index over the period 2001 through 2015. We find a much stronger negative correlation (where $r = -0.75, R^2 = 0.56$) between the two variables than Mitchell [101] suggested could be found (note that there was no improvement in the model fit by lagging the explanatory variable to 1, 3 or 6 months). By our estimates, U.S. dollar depreciation could have accounted for more than 50% of the rise in food prices. Naturally, this effect only explains food prices rises in Eurozone countries[1] importing food-commodities from the U.S. But these are the most widely reported figures concerning world trade.

[1] The Eurozone consists of Austria, Belgium, Cyprus, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Portugal, Slovakia, Slovenia, and Spain [42]

www.manaraa.com

6.4 The Role of Export Restrictions in the Crisis

As discussed above, the true origins of the crisis remain a matter of debate. However, evidence from many sources agrees that export bans and other trade restrictions put in place by large food-exporting countries, played a key role in aggravating the crisis. It is thought that while export restrictions did not cause the crisis, they did contribute to large price surges in at least three key food-commodity markets—wheat, rice, and maize [6, 102, 134, 26, 97, 74]. Countries that restricted or banned exports included several major rice exporters such as; China, Cambodia, Egypt, India, Indonesia, and Vietnam, several wheat exporters including; Ukraine, Kazakhstan, and Russia. Argentina and China were important maize exporters that also restricted exports [124]. Export restrictions were less important for other food-commodities, such as soybean and rape seed, as their export markets are heavily dominated by countries that did not imposed restrictions[1] [75]. The full list of export restrictions, shown in Appendix C1–3, was compiled from a variety of sources including; the World Trade Organisation's Trade Monitoring database [149], the Centre for Economic Policy Research's Global Trade Alert database [28], and from academic sources including: Shama [124], Slayton [127], and Dollive [38]. This list includes information on each trade restriction including: a start and end date; the type of policy enacted (tax, ban, or minimum price); and also which markets were affected.

Restricting exports directly impedes the physical flow of goods out of the country, thereby increasing the supply available to domestic markets [71]. This, at least theoretically, reduces the domestic price of food relative to the rising world price [97]. However, such actions have a ripple effect, passing on food price volatility to other exporters and make for more unstable international markets [2]. Countries on the receiving end of such policies face having fewer sources of supply from which to make their purchases. This invariably leads to higher prices. This 'tightening' of world markets can also heighten existing food security concerns, encouraging the hoarding of additional supplies [138]. The use of export bans can lead to a highly volatile situation which poses several interrelated risks to food import-dependant countries [118]. First, there is a risk that food-commodity prices will increase above affordable levels [118]. Second, the physical availability of food-commodities may be threatened. In extreme situations, several exporters may impose restrictions simultaneously, making it difficult for importers to secure alternative supplies in the time-frame needed [11]. While seemingly unlikely, this scenario has occurred in rice markets recently, when a flurry of restrictions in 2008 saw availability fall to the point where importers were reportedly unable to secure supply [12].

[1] Specifically, Australia, Canada, E.U. states, and the United States.

www.manaraa.com

The consequences of such an event is likely to be more severe for those countries that are heavily dependent on imports [147]. The third risk is that an international trading firm (the *counter-party*) will default on its contracts and fail to deliver any grain at all. This can often happen, as the firm may not necessarily own the goods that are to be delivered at the time the contract is agreed. If prices subsequently rise it may be better, financially, for the firm to default than trade at a loss in profits. In periods of rapidly rising food prices, counter-party risk leaves importers in the position of having to make more short-term purchases at higher prices [118].

Food price volatility and the associated risks outlined above are thought to have propagated throughout the international trading system as a cascading effect[1], with country after country banning exports in a struggle to keep their own domestic prices low [67]. However, by trying to insulating their own domestic market from international prices, each ban enforcing country simply forced the price volatility to be borne by other countries in the international market [102]. For example, Dollive [38] points to a clear contagion effect, between wheat export restrictions put in place by Ukraine, Russia, and Kazakhstan. Starting in July 2007, Ukraine announced strict export quotas that amounted to a complete ban on wheat exports [75]. Ukrainian exports fell approximately 77% the following year. As a consequence, many of Ukraine's largest trade partners switched to other major wheat exporters in the region, particularly, Kazakhstan and Russia. This resulted in greater pressure on Russian and Kazakh wheat exports, leading to the halving of wheat stocks in those countries. Worried by the effect of these developments on their own populations, who were struggling to afford staple foods, Russia and Kazakhstan both implemented their own export restraints in early 2008 [38].

This phenomena of cascading price volatility is similar to other economic shocks or 'perturbations' in that it can propagate globally via the network formed of countries connected together through trade [68]. Research has shown that a country's susceptibility to perturbations can be directly related to its level of integration in the trade network [87]. Similarly, Wu and Guclu [153] have theorised that clusters of interconnected nations in food-commodity markets could leave countries more or less robust to the impacts of sudden disruptions in other parts of the world. In both of these studies [87, 153], trade data has been analysed by modelling the world trade system as a network. In each market, multiple countries are connected together by many trade links. Taken together, these links give the network its structure (or topology). Export restrictions can affect this structure by removing specific trade links from the network.

[1] Giordani et al. [67] prefer to describe this phenomena as a 'multiplier effect', as a cascade implies a linear system of cause and effect, whereas export restrictions effect multiple countries simultaneously

www.manaraa.com

Each link removed in this way effects the network's overall topology, with some countries effecting the network more than others. The recent proliferation of export restrictions provides an opportunity to test these propositions and further explore the role the network structure of trade plays in propagating food-security risks in international food-commodity markets.

6.5 Methods

6.5.1 Modelling Export Restrictions in Networked Markets

Similar to the methods used in Chapter IV and Chapter V, we use international commodity trade data from the UN Comtrade database [141], this time selecting transactions over the period 2007–2012, for each of the three major international grain markets; wheat, rice, and maize[1]. These data are used to construct a weighted-directed adjacency matrix of a network, where vertices represent countries, and edges denote the volume (measured in US$ worth) of imports/exports of a commodity between those countries. The Matlab function to do this is reproduced in Appendix C4. Note that the Comtrade data used are of a low resolution, with changes in trade being recorded only on a yearly basis up to 2010, then monthly [141]. The monthly Comtrade data was not extensive enough to cover the period of this study. As such, the effects of short-term export restrictions (i.e., those in effect < 12 months) may not be fully accounted for in the results.

We model these data to account for dynamic changes in the network topology, such that all measures of the network act as a function of time. Here, time is denoted as $t_y = 2007 + y$, and $y = 0, 1, 2, ..., t$, representing the years 2007-2012. we arranged the Comtrade data for each year (t_y) in a $n \times m$ adjacency matrix $\mathbf{A}(t_y)$, where elements $a_{ij}(t_y)$ of $\mathbf{A}(t_y)$ indicate yearly trade between the $n = 1, 2, ..., i$ food-exporting countries and the $m = 1, 2, ..., j$ food-importing countries. Furthermore, for every non-zero element in the adjacency matrix, the volume of trade between exporter i and importer j, recalculated as a percentage of country j's imports, is recorded in a similar $n \times m$ weight matrix $\mathbf{W}(t_y)$.

The purpose of this method is to measure how export restrictions effect the structure of the network. From a topological perspective export restrictions act by removing edges from the network. We can measure this process via the edge-detachment rate D_k, i.e., the rate at

[1] HS-02 classification system was used to select the product codes; 1001, 1006, and 1005 denoting trade in wheat, rice, and maize respectively.

www.manaraa.com

which edges are removed from vertices of degree k. More specifically, to measure a single case d, where a trade restriction imposed by exporter i affected trade with the importer j in year y, leading to the removal of the link $a_{ij}(t_y)$, in the network's adjacency matrix, we can compute

$$d_{ij}(t_y) = \begin{cases} 1, & \text{if } a_{ij}(t_{y-1}) - a_{ij}(t_y) = +1 \\ 0, & \text{otherwise.} \end{cases} \tag{6.4}$$

Thus $d_{ij}(t_y) = 1$ signifies a complete loss of trade between i and j in year y. The frequency of these events is simply the sum of all instances of d in a given year. To understand how the likelihood of export restrictions vary with a country's position in the network, we apply a sorting algorithm so that countries are arranged into subsets, with each subset relating only to importing countries of k degree. Taking the average number of edge-removals in each subset thus gives the edge-detachment rate D_k

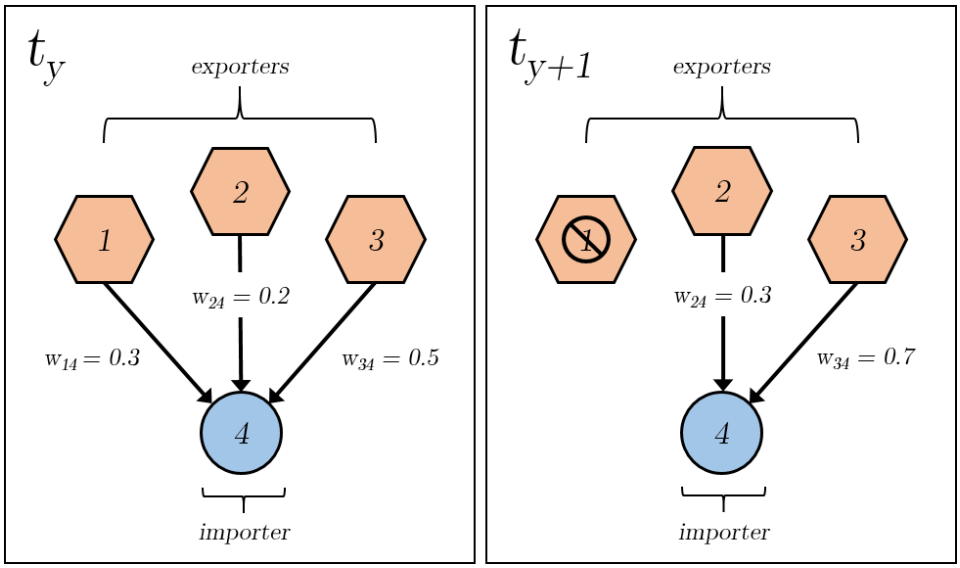


Fig. 6.12 Schematic of a single export restriction in a simple network.

We illustrate this process in figure 6.12, showing the result of a single export restriction in schematic form. In our example, at time period t_y, exporters 1 through 3 are trading with importer 4. Supply is measured by the trade link weights, so that where $w_{14}(t_y) = 0.3$ denotes that country 1 supplies 30% of country 4's imports. In the following year t_{y+1}, the government of country 1 places a restriction on exports, leading to the loss of trade between 1 and 4. In this instance, we would record $d_{14}(t_{y+1}) = 1$ satisfying the condition of equation (6.4). The percentage of trade lost due to this export restriction is calculated by multiplying $d_{14}(t_{y+1})$

www.manaraa.com

by the trade link weight in the previous year $w_{14}(t_y)$. Note that in the following year the link weights are updated so that $\sum_i w_{i4}(t_{y+1}) = 1$, to allow consistent comparison across years. We use percentages here so that the edge weights show which trading partners are important locally, to each specific importer regardless of size and trading volumes (which can vary greatly across markets). However, the trade values can also be calculated in terms of US$ of trade if required.

6.6 Results

Before looking at the impact of export restrictions, we report summary network statistics for the three food-commodity markets in question. First, we calculate the number of trade links to-or-from a country in a given year (i.e., the vertex degree k) as follows. For an exporting country (out-degree),

$$k_i(t_y) = \sum_{j=1}^{n} a_{ij}(t_y), \tag{6.5}$$

and for an importing country (in-degree),

$$k_j(t_y) = \sum_{i=1}^{m} a_{ij}(t_y). \tag{6.6}$$

The average number of trade links per country is denoted as $\langle k \rangle$. The relative frequency by which countries are connected by these degrees is expressed by the network's degree distribution $\mathbb{P}(k)$. The cumulative degree distribution (i.e., $\mathbb{P}(K \geq k)$) for imports and exports in each of the three markets are shown in figure 6.13. We have covered the methods for studying the precise functional form of these distributions at length in Chapter IV, so we will not reiterate here. However, it is worth noting that the rice and maize markets, which were not studied in the previous chapter, follow much the same distribution as other markets. That is to say that the probability of finding other exporters with a similar degree, decays in a highly non-linear trend.

Like many other real-world networks, these markets are sparse [13], that is to say, the average number of trade links is far less that the total number of exporting countries in the network, so that $\langle k \rangle \ll n$. In the case of the wheat market, $\langle k \rangle = 11$, whereas $n = 187$. In addition, the degree range is between 1–110 (with the United States, the world's largest wheat exporter, having the highest degree), whereas the mean degree is only $\langle k \rangle = 11$. Similar trends are found for both rice and maize markets as shown in Table 6.1. As such, we can say the vast majority of exporters tend to trade with relatively few countries. This is perhaps reflecting the fact that trade with geographically neighbouring countries (i.e., cross-boarder trade) represents

www.manaraa.com

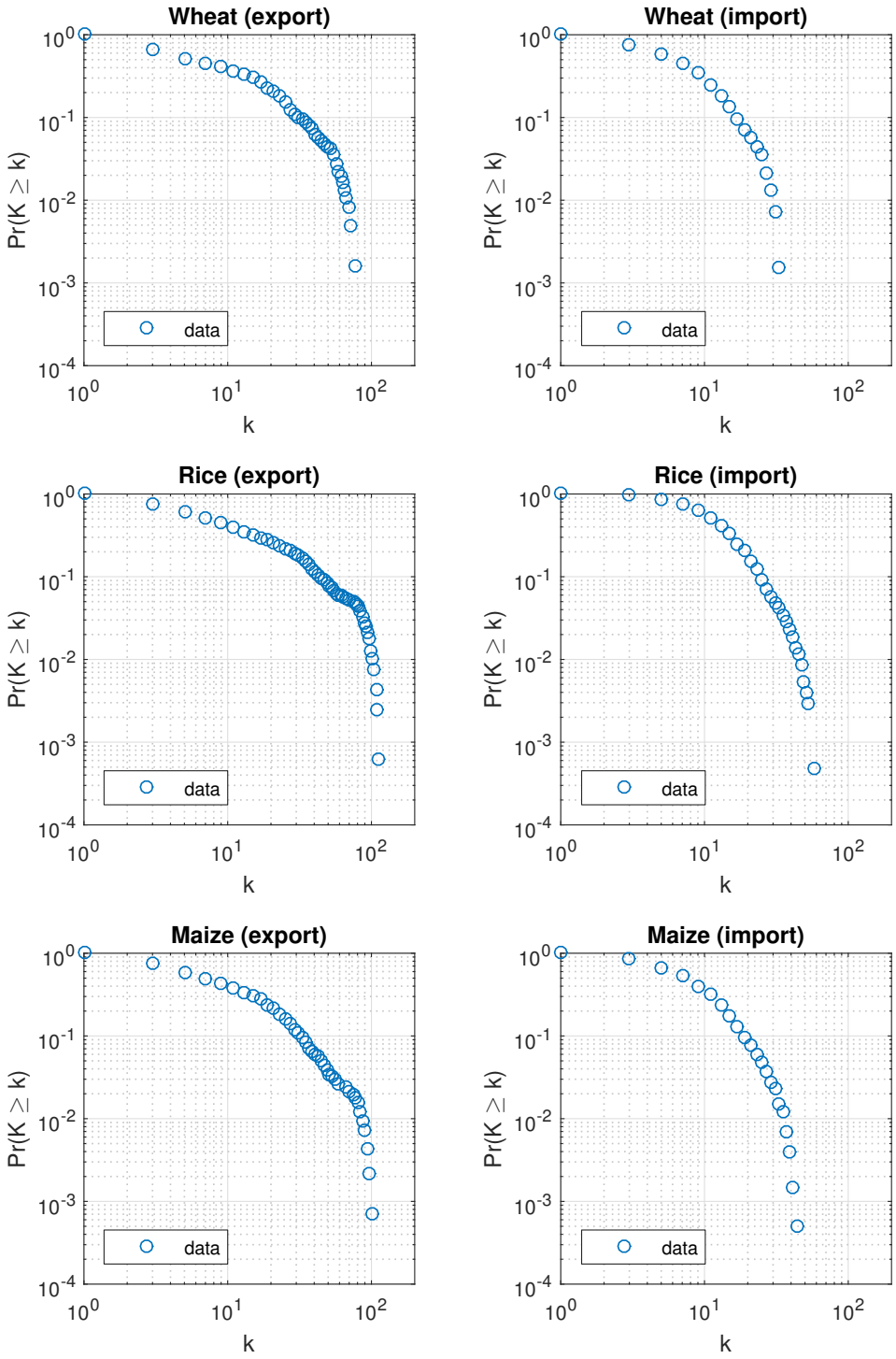


Fig. 6.13 Import and export degree distributions of selected markets.

www.manaraa.com

a large portion of trade globally [141]. Only a handful of exporters have an extensive number of trade links, but those who do, trade with almost every other country active in the market. For example, the top four exporting countries in the global wheat market are the USA ($k = 110$), Canada ($k = 76$), France ($k = 74$), and Australia ($k = 47$). A measure of export degrees gives us some indication as to the maximum number of food-importers that could be affected by a trade restriction imposed by a k degree exporting country. For example, a trade restriction imposed by the United States, with its 100+ country trade network could be severely damaging to world wheat supplies. If we compare that to Ukraine, who only supplied wheat to 38 other countries during the time under study, we would be looking at under half the network affected. Of course, this does not tell us the likelihood of either country banning wheat exports, only the theoretical outcomes of such a policy decision.

Table 6.1 Network statistics for selected commodity markets.

		Wheat	Rice	Maize
Export				
Mean node degree	$\langle k \rangle$	11	16	13
Degree Range	k_{min}, k_{max}	1–110	1–147	1–138
Import				
Mean node degree	$\langle k \rangle$	9	16	11
Degree Range	k_{min}, k_{max}	1–38	1–58	1–42
Total number of links	$\|E\|$	1398	2697	1767

We can also see that the total number of trade links in each market $|E|$, is positively correlated with higher values of $\langle k \rangle$. Similarly, having more trade links leads to a wider degree variance, giving a larger range of degrees. Of the commodity markets studied, the market with the most trade links is rice, (2697 links), more than twice that of wheat (1389 links), and significantly more than maize (1767 links). This is surprising, given that rice is often called a 'thinly' traded commodity (as 90 % of all rice is consumed domestically [75]), but from a network perspective, rice is the most actively traded commodity of the three. Similarly, the rice market also has the highest mean degree $\langle k \rangle = 16$, and the largest range of k degree nodes (1–58).

www.manaraa.com

6.6.1 The Impact of Export Restrictions

In real-world trade networks, importers are connected to multiple exporters, and vice versa, so while a trade restriction may affect many parts of the network, it can also be mitigated to some degree by having a redundancy in supply. It follows that countries with a higher degree (i.e., those with many trade links), would have a greater degree of redundancy. Conversely, countries that depend on a limited number of suppliers are theoretically more vulnerable to disruptions [153]. However, we must consider whether being more integrated into the network also increases a country's exposure to these risks. That is to say, is there a relationship between country's position in the network and the likelihood of being affected by export restrictions?

Although individual export restrictions are impossible to predict and do not occur in a strictly regular pattern, we find a very general statistical regularity to these events. Figure 6.14 shows a histogram of the yearly rate of trade links being removed from the network, measured per k-degree. The data fits a linear model, with $r = 0.75$ (with a goodness of fit $R^2 = 0.56$). That is to say, that high k countries are affected more often than low k countries, and that this relationship is monotonically linear in k. Specifically, given a country with degree k and another with a degree of $2k$, we would expect the latter to be approximately 1.4 times more likely to be affected by an export restriction. This suggests, while each individual trade link in the network should have an equal chance of being affected, this process has a form of preferential detachment when measured by a country's degree. From these data, we can also say that the edge-detachment rate $D_k \sim k + \alpha$, where α is a linear scaling coefficient. There is some difficulty in estimating D_k for high degree countries, as there are so few observations in the data (due to their relative scarcity in the network), but their seems little reason to suggest this trend would not hold.

In market networks, not all trade links are of equal value. The volume of trade through each link can vary significantly, even between those connected to the same importer. To take this variation into account, we estimate the expected losses to supply from the effect of an export restriction. Figure 6.15 shows the edge weights (measured as a percentage of total imports) for k degree importers in the wheat market over the period 2000-2012. Naturally, the link weight for importers with only a single trade link is always $w_1 = 1$, as that sole edge must supply 100% of imports. We can see that where $k > 1$, there is a wide range of edge weights. Interestingly, even some high degree countries source 90-99% of their imports from just a single supplier, despite trading with many others. However, if we look at the average edge weights $\langle w \rangle_k$, importers tend to source their supply equally from each of their trade partners.

www.manaraa.com

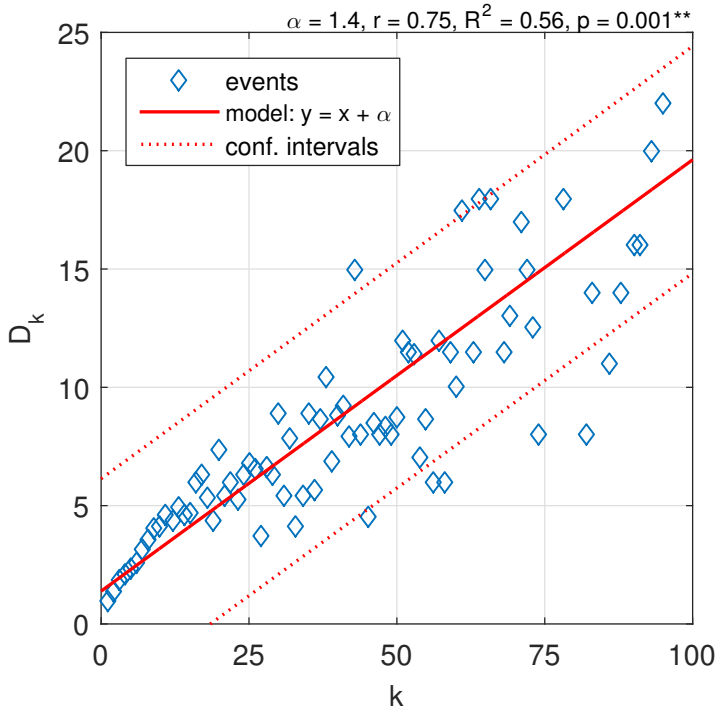


Fig. 6.14 Number of edges removed from the wheat market's network on an annual basis.

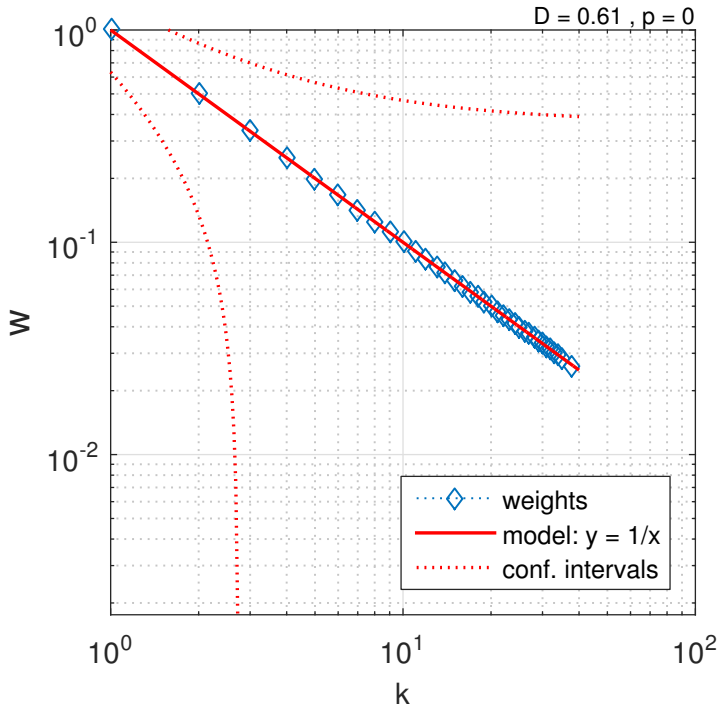


Fig. 6.15 Edge weights measured against importer degree, k.

www.manaraa.com

As such, the mean edge weight distribution can be modelled simply as $\langle w \rangle_k \sim 1/k$. This means that in high degree countries, approximately 90% of trade links supply no more than 10% of a particular country's imports, with many edges supplying less than 1%. This places a limit on the potential impact of a single export restriction to these high degree countries.

6.6.2 Developing a Composite Risk Index

Now that we have some insight on how trade networks are structured, we want to understand whether the benefits of being more integrated into the network (i.e., increased redundancy) outweigh the risks, through increased exposure to export restrictions? To do this, we combine both factors outlined above into a composite risk index [94], which typically follow the form

risk index = impact of risk event $\times$ probability of occurrence

To measure the impact of a risk event, we calculate the expected losses to k-degree countries from a single export restriction $\mathbb{P}(w_k)$. This value gives an indication of how much redundancy a country is likely to have, in terms of the effect on import supply (with lower values being better). For the probability of occurrence, we take the probability distribution function for one edge being removed from the network $\mathbb{P}(d_k = 1)$. This value quantifies the increased exposure to export restrictions in higher degree countries. The composite risk index then combines these values as follows,

$$R_k = \mathbb{P}(w_k) \cdot \mathbb{P}(d_k = 1), \tag{6.7}$$

where the risk index R_k is a unit-less value on a scale $[0, 1]$ that allows us to compare risk across the network. By this measure, values of $R_k \gg 0$ implies that k degree countries are at greater risk from export restrictions. By design, values of R_k display *positive homogeneity* [94], that is to say, if the value of R_1 is twice as large as R_2, the former faces twice the risk.

Figure 6.16 shows the completed composite risk index for wheat importing countries. We find that the level of risk is highest for those countries with low degrees (particularly where $k < 5$). The risk index then decays for increasing values of k, with risk levelling off at high values of k. While low degree countries are unlikely to be affected by export restrictions due to their low connectivity, they tend to be highly or entirely reliant on just one trade partner to supply the majority of their imports. As such, a disruption to supply could potentially lead to these importers being cut off from international markets altogether unless alternative trade contracts can be negotiated. Countries with very high degrees have the lowest levels

www.manaraa.com

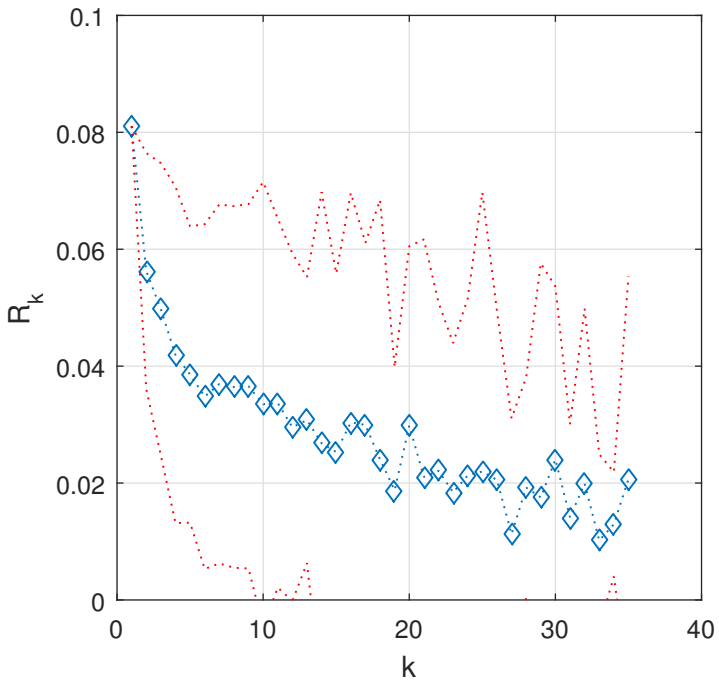


Fig. 6.16 Composite risk index for the wheat import market. Shows the risk R_k to importers of degree k, from a single export restriction.

of risk, despite being more likely to be affected by export restrictions. In effect, high degree food-importing countries are able to diversify their supply such that a disruption to one trade link will have only a minimal impact on the overall flow of imports.

6.7 Discussion

6.7.1 A Policy for Strategic Reserves for Import Dependent Countries

Following the recent spate of export restrictions, numerous proposals have been put forward in order to mitigate both import supply and price risks. These proposals include both bilateral or multilateral agreements to limit export restrictions, increasing physical reserves of grain in both the public and private sectors, creating so-called 'virtual' reserves at regional and international levels, and various improvements in coordination, logistics, and to the flow of information across nations (see Torero and von Braun [136] for details). A natural response to volatility in international markets is to refocus on national self-sufficiency in order to provide a sense of food security. While self-sufficiency reduces a country's reliance on trade, it comes with its own risks and costs [96]. For example, using limited land and resources to grow low value staple crops may have a high opportunity cost compared to using those same resources to produce

www.manaraa.com

high value goods for export. In addition, encouraging agriculture in increasingly marginal areas can leave production more susceptible to weather related shortfalls and climatic risks [90].

The policy we focus on in this study is the practice of maintaining strategic grain reserves. The stockpiling of grain reserves has a long history in food security debates. After the recent crisis, many countries have renewed their interest in such policies, particularly in the heavily import-dependent Middle East and North Africa (MENA) region [118]. Since 2011, these countries have been planning to increase their strategic storage capacity to accommodate reserves that will last up to one or two years in some cases [90]. Wright [152] argues that this policy response is almost inevitable given how unreliable suppliers have proven to be in a crisis. Pursuing a policy of building strategic reserves creates two challenges that need to be addressed: the determination of optimum stock levels, and the decision on how to finance the associated cost of maintenance [136]. In terms of cost, physical reserves are expensive and must be rotated regularly. Unfortunately, those heavily import-dependant countries that are likely to need extensive reserves, are generally those least able to afford it [136]. There are several additional costs that may need to be considered [90]; capital investment is needed to expand capacity by building new storage facilities; transport infrastructure may require upgrading; there are costs for training staff in how to manage these reserves; and countries must be able to afford the initial increase in grain purchases require to fill reserves. Maintaining a transparent and accountable governance for these reserves could also be considered a cost [90]. Once built, strategic reserves can be used to mitigate intermittent price risks and smooth out market volatility by selling stock to domestic consumers at lower than market prices. However, such a strategy offers no protection against structural, long-term price increases [90].

Strategic reserves can be used to manage supply risks, i.e., the risk that grains will not be available to purchase from international markets. The strategy is relatively simple in theory, in that a government only needs to ensure that enough grains have been stockpiled to supply the domestic market through a period of unavailability [90]. Small reserves can also give the importing countries the time they need to secure alternative supplies and avoid outbreaks of panic buying and hoarding [104]. Sadler and Magnan [118] suggests that calculating optimal levels might be formulated as follows: the average per capita consumption by month, multiplied by the number of citizens, and by the number of months a stockpile should cover. The first two variables are relatively easy to calculate from government census data and food surveys. The last variable, however, is dependent on a further set of factors including a country's level of import dependency, risk tolerance, and, importantly to this study, a country's vulnerability to

www.manaraa.com

supply disruptions [146].

By understanding which factors make a country more or less vulnerable to supply disruptions, such as export restrictions, we can shed some light onto which set of countries should maintain the highest levels of strategic reserves. In the case of the wheat commodity market we can see from the risk index in Figure 6.16, that low degree importers (where $k < 10$) face the highest level of risk. Here, each export restriction is likely to affect a significant percentage of supply (ranging from 10% to 100% of imports per link, see Figure 6.15). Those countries with the highest vulnerability to import supply should benefit the most from maintaining higher stockpiles and this should justify the cost. However, while these countries have the highest expected losses, they are not expected to be affected very often. This could work in favour of a strategic stockpile policy, as maintaining domestic stockpiles would not be viable in a climate of high frequency supply disruptions, as having to replenish stocks on a regular basis could rapidly become unaffordable to governments pursuing such a policy. High degree importers (where $k > 10$) face a relatively higher rate of supply disruptions, but due to their increasing redundancy of supply can mitigate the consequences of each restriction far better. Increasing levels of supply redundancy not only lowers their level of risk, but also leads to lower expected losses, as each link generally accounts for only a small proportion of overall imports. These countries can in effect, leverage their position in the network to mitigate the supply risk. This is an example of a 'risk-sharing' approach to supply chain resilience [76], where importers seek as many suppliers as possible to be able to compensate for the loss of one supplier.

To illustrate how these metrics could be applied in developing a policy of strategic reserves, we apply our method to Yemen as a test case. Yemen is a net-food-importing country that is dependent on international trade for up to 75% of its domestic food consumption [7]. For wheat, one of the country's staple foods, dependency on trade is closer to 90%. In addition, 96% of the population are net-food buyers [113], and Yemeni households in the lowest decile of income have to spend up to 55% of their monthly expenditure on food [145]. As such, the food security situation in Yemen is extremely sensitive to events in international markets [7]. Yemen is the 18th largest wheat importer globally [149] and imports from approximately 16 countries on average (over the years 2000-2012), so the mean degree $\langle k \rangle = 16$ (rounded to the nearest integer). Total imports are approximately US$1,000,000,000 of wheat per year. An average trade link to Yemen supplies approximately $1/k$ of imports per year, so that the mean import link-weight $\langle w \rangle \sim 0.0625$, or in US$62,500,000. The expected number of links effected by export restrictions per year $D_{16} \sim 3$. As such, we can calculate the yearly expected

www.manaraa.com

loss to import supply as $0.0625 \times 3 = 0.1875$, or US$187,500,000, approximately 0.19% of Yemen's yearly food import bill. These values could be used as an estimated budget for Yemen food security programme to hedge against the risk posed by export restrictions in the wheat market.

6.8 Conclusion

In this chapter, we built on previous research which has shown that a country's susceptibility to economic shocks is directly related to its level of integration into the trade network's structure. Informed by a network theory approach, we analysed three major food-commodity markets (wheat, rice, and maize) using UN Comtrade data. We model each commodity market as a network to considered whether a country's position in this network affects the likelihood of it being affected by an export restriction. We find that the average yearly rate at which food-importing countries are affected increases as a function of a country's degree. It follows that the more connected a country is in the network, the more likely it is to be on the receiving end of an export restriction. However, we also find that individual trade links to high degree countries tend to supply only a small fraction of overall import volume. As such, removing one of these links is unlikely to result in significant shortfall in aggregate supply. We combine these two factors into a composite risk index to compare the risk posed by export restrictions to each country based on its position in the network. We interpret these risks in terms of developing a policy of strategic reserves and provide a case study example showing how these network metrics could be applied to improve a country's food security situation.

It should be noted that food security is a multi-faceted problem, whereas we focus on a narrow dimension of the problem. Before making policy recommendations on a per country basis, we would need to take into account several other important variables. In particular, a country's import dependency ratio (i.e., the ratio of net imports to domestic consumption) and foreign exchange reserves play an important role in determining how much capacity a country has to deal with both price and supply risks (including export restrictions). The import dependency ratio is indicative of a country's dependency on foreign imports to satisfy domestic demand [78]. This is important in terms of determining how much of a shock to international markets will be passed on to domestic consumers (usually in terms of food price rises). The value of a country's food imports as a share of foreign exchange reserves is also important, as it determines how easily they can pay for imports and adjust to higher prices. Low levels of foreign exchange reserves are particularly a problem for many developing countries [146].

www.manaraa.com

While these variables are undoubtedly an important factor in a country's overall food security outlook, they are largely predicated on a country's own geography and economic situation, rather than its position in the network, so as such they are beyond the scope of this study.

www.manaraa.com

Chapter 7

Conclusion and Outlook

7.1 Research Contributions

The broadest contribution of this thesis is in the exploration of the intersection between economic theories of trade and applications in graph theory. The rise of globalisation and the increasing importance of trade in the global economy shows that work in this area is well motivated. While the field is relatively new, with only a small range of existing literature, we have shown that these two, seeming disparate, fields can be effectively brought together. More specifically, we have shown through a variety of models, experiments, and simulations that real-world policy issues can be addressed using a quantitative network approach. This approach is important for two reasons. First, it will help the policy community understand how the structure of trade networks affects the behaviour and dynamics of international commodity markets, and second, by developing policies that favour more stable and equitable network structures, we can aid in guiding the restructuring of future trade networks.

In the remainder of this concluding chapter, we provide an overview of the thesis results from each of the three projects in Section 7.2. We outline areas for further development in Section 7.3, and in Section 7.4, we offer some discussion on the policy implications of our findings with regard to developing international trade policy.

www.manaraa.com

7.2 Overview of Thesis Results

7.2.1 Chapter IV: Growth in Network Models of Commodity-Trade

In Chapter IV, we found that reports of power-law behaviour in the degree distributions of international trade networks are not as categorical as the literature would imply. We outlined the common critiques of these works and identified a wide-range of empirical degree distributions using methods that are more appropriate. Having found that true power-law degree distributions are unlikely, we re-evaluated the use of the so-called 'preferential attachment' model that is widely interpreted as the underlying model of network growth in the trade literature. Using alternative methods, we showed that connectivity growth in many trade networks is in fact *constrained*, due to the limited opportunities for new countries (vertices) to enter these markets, resulting in far different network topologies than the preferential attachment model predicted. Taking these findings into account, we developed a new single parameter model that reproduces the structural features of a wide array of international commodity trade networks. An interesting feature of this model is that it can resolve to the Barabási-Albert (B-A) preferential growth model [14], or the Erdős-Rényi (E-R) random graph model [41], simply by tuning the model's only parameter, giving us the ability to compare the empirical trade networks to both these well-known theoretical ideals. We concluded by discussing the wide reaching implications of our findings. First, in applying the model to estimate future market growth, and second, in terms of understanding of the structural robustness of critical commodity markets. In the first case, we found that the B-A preferential attachment model overestimated the growth in network size and connectivity by more than two orders of magnitude when compared to the constrained growth model with empirically estimated parameters. In the second case, we found that our model implies that real-world trade networks may be far less robust to error type perturbations than the B-A model would suggest.

7.2.2 Chapter V: Price Variation in Network Models of Commodity-Trade

In Chapter V, we develop further the graphical exchange model. We applied this model to data from international commodity markets in order to measure the effect of the network structure on price asymmetries. Using the model, we investigated the local determinants of price variation, and measured how global prices vary between networks with different topological structures. In the first case, we find that increasing an exporter's degree (i.e., the number of importers that a country trades with) also increases its prospects of obtaining better prices. However, we also find that increased competition in the market, measured by the average nearest neighbours'

www.manaraa.com

degree, rapidly reduces the maximum price an exporter can command. In the second case, we simulated several synthetic markets using the network growth algorithm developed in Chapter IV, so that we could measure global price inequality (measured here by the Gini coefficient) across networks with a range of different degree distributions. Our results suggest that price inequality increases as network degree distributions become more heterogeneous. Along this scale, we find most real-world export markets are relatively equal. However, these markets are also found to be in a very sensitive position, so that any change in the network's rate of growth could dramatically increase price inequality between exporters in these markets.

7.2.3 Chapter VI: The Risk of Export Restrictions to Food-Commodities

In Chapter VI, we explored the causes and consequences of the global food crisis—a recent episode of extreme food-price volatility that deeply affected many consumers in food-import dependant countries. One of the key factors that exacerbated the crisis were export restrictions placed on trade in several major grain markets. We found that an importing country's susceptibility to export restrictions and similar shocks can be directly related to the structure of the international trading network. To investigate this issue further, we developed a network theory approach to measure the effect of export restrictions in three of the major food-commodity markets. Our findings showed that the average yearly rate at which food-importing countries are affected by export restrictions increases as a function of degree. However, we also found that trade links to high degree importers tend to supply only a small fraction of the importer's overall supply volume. As such, removing one of these links is unlikely to result in significant shortfall in supply. We used this analysis to develop a composite risk index that assesses the risk posed by export restrictions to food-importing countries, based on their position in the network. The index shows that the level of risk falls significantly for high degree countries despite the aforementioned exposure to export restrictions. We used this risk index to develop a policy of strategic reserves, and provided a case study example, showing how these network metrics could be applied to improve a country's food security situation.

7.3 Future Work

This section outlines future research that could directly extend this study and improve our ability to quantifying commodity markets in terms of a network topology, and develop new tools for investigating real-world policy concerns. We see the need for future work in four

www.manaraa.com

specific areas, the first three relate to the constrained growth model developed in Chapter IV, and the last relates to the risk index developed in Chapter VI.

1. Using the constrained growth model developed in Chapter IV, we might also consider investigating correlations between country degree and second-order topological properties, such as the average nearest neighbour degree, and local clustering coefficients. There are already many reported values in the literature that could be used to further validate our model.

2. We could also work on validating and improving the constrained growth model's performance in predicting future market growth by back-testing predictions against historical data.

3. Based on our conclusion to Chapter IV, work on quantifying the robustness of networks generated by the constrained growth model would be another area to consider.

4. In Chapter VI, we found that, generally speaking, higher degree countries are less at risk for export restrictions than low degree countries. However, an important caveat to these results is that each set of k degree countries has its own distribution of risks that changes qualitatively between sets. Therefore, we could improve the accuracy of our risk assessment in this area by considering these changes. Further work in this area could apply a Bayesian approach to measure the conditional probability of export restrictions based on country degree [83].

7.4 Policy Implications

In this final section we discuss the implications of our findings with regard to developing international trade policy. Methods for improving growth (both in terms of network size and connectivity) in international commodity markets is increasingly important to policy makers, as international trade itself, has become the primary means of stimulating economic development in many of the world's least-developed countries [126]. This is also a symbiotic relationship, as furthering the development of export-orientated industries aids in integrating these countries into the global economy [139]. However, as noted in the introduction to this thesis (Chapter I), the use of direct market interventions tends to influence outcomes in favour one set of consumers (typically in developed industrial nations) over another (in developing and emerging economies) [131]. Therefore, we need to think about a more systems level approach that can work for all involved. One way of achieving growth that benefits both sides is through the use

www.manaraa.com

of *Foreign Direct Investments* (FDI), which the World Trade Organisation (WTO) define as investments made by businesses in a developed country (referred to as the 'home' country) to acquire assets in a developing country (the 'host' country) with the intent to manage those assets (as opposed to passive investments via the purchase of company stock) [22]. Foreign direct investment is usually achieved through mergers and acquisitions, or through the creation of new facilities such as ports, mining operations, transport infrastructure, logistical services, and export-orientated agriculture. The WTO [22] have argued that FDI is the only way to reverse the growing inequalities between the world's least developed countries and the rest of the global economy. Where FDI differs from passive investment and foreign aid, is that it also brings with it key resources that are lacking throughout the developed world, including financial capital, the transfer of new technologies, and the organisational, managerial, and marketing skills of foreign companies [23]. Foreign direct investments can also improve the 'home' company's competitive position by reducing its labour costs, transactions costs, and the cost of inputs, which is why it has become the favoured policy instrument of the WTO and many developed economies [22]. Foreign direct investment has had some dramatic successes, for example, China's rapid development heavily relied on FDI, which allowed the Chinese government to restructure and diversify its economy, making it more competitive in global markets [24].

Here, we are particularly interested in FDI for its ability to promote growth in exports from developing countries. We know from empirical studies, that foreign owned companies (acquired through FDI) tend to export far more than their domestic counterparts [23], and it is thought that these companies tend to have greater knowledge of global markets and so naturally have a comparative advantage in international trade. These advantages are then thought to 'spillover' into the host country and improve the ability and propensity of domestic companies to also export [23]. To encourage FDI several policy instruments are available, such as targeted subsidies, the establishment of so-called 'special economic zones' in the host country (i.e., an area in which trade laws may differ from the rest of the country), and in developing bilateral/regional trade deals [22]. The latter of these policy instruments has been used to some success, most notably with the creation of the European Economic Community (EEC, now the European Union), but also with NAFTA (the North American Free Trade Agreement), ASEAN (Association of Southeast Asian Nations), APEC (Asia-Pacific Economic Cooperation), and through OECD (Organisation for Economic Co-operation and Development) negotiations. These initiatives work in two ways, first by removing internal barriers to trade, and second, in that they provide the security of future market access that international companies

www.manaraa.com

value [22]. Trade deals have shown to have a positive impact on FDI flows. For example, the European Union has stimulated substantial investment, between countries in the Union and from countries outside the single market [79]. However, concerns have been raised that these initiatives were not open to many of the world's least-developed countries, thereby, only added to the growing inequalities these countries face [24, 23]. If these policy instruments were to be used to promote global network growth in commodity markets, these failings would need to be addressed.

7.4.1 Application of Thesis Results

Having outlined the policy space on promoting growth in international commodity markets, we now turn to consider the implications of our network modelling on these policy decisions. The network model we developed in Chapter IV shows that growth is primarily constrained due to issues with market access, so if our aim is to improve growth, we should choose those FDI development policies that are most likely to address this specific issue. As noted above, many existing FDI initiatives are not open to those least-developed countries that would benefit the most. However, we should also be aware that faster growing networks could result in far different topologies (i.e., moving towards the scale-free structure of the B-A model) than what we find in existing empirical networks, and this could have both costs and benefits in other policy areas. For example, our work in Chapter V showed that while current network structures allow for relatively equitable price formation (measured by the Gini coefficient), small changes to the growth model leads to significant increases in price inequality. In addition, if we consider the implications of our work in Chapter VI, similar changes to topology would lead to an overall increase in network connectivity. This is important, as it would allow for more diversification of supply, and a likely reduction in the risk posed by export restrictions to importers across many commodity markets. However, these benefits are unlikely to be shared equally, as a move towards a scale-free network structure results in highly heterogeneous degree distributions. Another dimension to consider, is that the least-developed countries tend to trade in only a handful of commodities internationally: as such, improving *in*-market diversification (i.e., increasing the number of suppliers of a given commodity) would benefit those countries more than *between*-market diversification, which does more to benefit the multi-national companies based in the developed world. Taken together, these results show that any policy that aims to improve market growth will always be a trade-off between price equality, supply risk, and economic development.

www.manaraa.com

References

[1] Abbott, P. (2009). Development dimensions of high food prices. Working paper 18, OECD Publishing.

[2] Abbott, P., Hurt, C., and Tyner, W. (2008). Whats driving food prices. Technical report, Issue Report, Farm Foundation, IL.

[3] Agricultural Food and Policy Center (2008). The effects of ethanol on texas food and feed. College Station, TX: Texas A&M University.

[4] Akyıldız, E. (2010). *Proten Domain Networks: Anaylsis of Attack Tolerance*. PhD thesis, Middle East Technical University, METU.

[5] Albert, R., Jeong, H., and Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382.

[6] Anderson, K. and Nelgen, S. (2012). Trade Barrier Volatility and Agricultural Price Stabilization. *World Development*, 40(1):36–48.

[7] Anwar, J. (2013). Yemen - Market Study. National Yemen. Online: www.nationalyemen.com/2013/03/03/wheat-imports-cause-yemen-heavy-losses/. Accessed: 2015-10-13.

[8] Arrow, K. J. and Debreu, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica: Journal of the Econometric Society*, pages 265–290.

[9] Axtell, R. (2005). The complexity of exchange. *The Economic Journal*, 115(504):F193–F210.

[10] Baffes, J. (2007). Oil spills on other commodities. *Resources Policy*, 32(3):126–134.

[11] Bailey, R. and Willoughby, R. (2013). Edible Oil: Food Security in the Gulf. Briefing Paper November, Chatham House, London.

[12] Baker, A. (2012). Desert Dreams : Can the Middle Eastern Country of Qatar Learn to Feed Itself ? *Time*.

[13] Barabási, A.-L. (2012). *Network Science*. Northeastern University.

[14] Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):11.

[15] Barabási, A.-L., Albert, R., and Jeong, H. (1999). Mean-field theory for scale-free random networks. *Physica A*, 272(1):173–187.

المنارة للاستشارات

www.manaraa.com

[16] Barabási, A.-L. and Bonabeau, E. (2003). Scale-free networks. *Scientific American*, 288(5):50–59.

[17] Barrat, A., Barthelemy, M., Pastor-Satorras, R., and Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–3752.

[18] Bascompte, J. and Jordano, P. (2007). Plant-Animal Mutualistic Networks: The Architecture of Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 38(1):567–593.

[19] Baskaran, T., Blöchl, F., Brück, T., and Theis, F. J. (2011). The heckscher–ohlin model and the network structure of international trade. *International Review of Economics & Finance*, 20(2):135–145.

[20] Besanko, D. and Braeutigam, R. (2008). *Microeconomics: An integrated approach.* NJ: Wiley.

[21] Bhattacharya, K., Mukherjee, G., Saramäki, J., Kaski, K., and Manna, S. S. (2008). The international trade network: weighted network analysis and modelling. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(02):P02002.

[22] Blackhurst, R. and Otten, A. (1996). Trade and foreign direct investment. Technical report, World Trade Organisation, Geneva.

[23] Blomström, M., Lipsey, R. E., and Zejan, M. (1992). What explains developing country growth? *NBER Working Paper*, No. 4132, August.

[24] Borensztein, E., De Gregorio, J., and Lee, J.-W. (1995). How does foreign direct investment affect economic growth? *NBER Working Paper*, No. 4132, August.

[25] Borgatti, S. (2015). *Graph Theory*. Analytic Technologies, Lexington, KY.

[26] Bouët, A. and Laborde, D. (2011). Food crisis and export taxation: the cost of non-cooperative trade policies. *Review of World Economics*, 148(1):209–233.

[27] Buchanan, S. (1987). *CRB commodity year book*. Commodity Research Bureau, Chicago, IL.

[28] Centre for Economic Policy Research (2009). Global Trade Alert Report. *Interpreting*, (June):1–10.

[29] Chakraborty, A. and Manna, S. (2010). Weighted trade network in a model of preferential bipartite transactions. *Physical Review E*, 81(1):016111.

[30] Chu-Carroll, M. (2007). Maximum flow and minimum cut. Online: goodmath.scientopia.org/2007/08/07/maximum-flow-and-minimum-cut/. Accessed: 2015-03-30.

[31] Clark, H. (2011). The real wealth of nations: Lessons from the Human Development Report. Technical Report Development Co-operation Report: 50th Anniversary Edition, OECD, Washington, DC.

[32] Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-Law Distributions in Empirical Data. *SIAM Review*, 51:661–703.

www.manaraa.com

[33] Clauset, A., Young, M., and Gleditsch, K. S. (2007). On the frequency of severe terrorist events. *Journal of Conflict Resolution*, 51(1):58–87.

[34] Cohen, R., Erez, K., Ben-Avraham, D., and Havlin, S. (2001). Breakdown of the internet under intentional attack. *Physical review letters*, 86(16):3682.

[35] Cohen, R. and Havlin, S. (2010). *Complex Networks: Structure, Robustness and Function*. Cambridge University Press, Cambridge, UK.

[36] Currarini, S., Marchiori, C., and Tavoni, A. (2014). Network economics and the environment: insights and perspectives. *Environmental and Resource Economics*, pages 1–31.

[37] Devanur, N. R., Papadimitriou, C. H., Saberi, A., and Vazirani, V. V. (2002). Market equilibrium via a primal-dual-type algorithm. In *null*, page 389. IEEE.

[38] Dollive, K. (2008). The impact of export restraints on rising grain prices. Technical report, U.S. International Trade Commission, Washington, DC.

[39] Drèze, J. and Stern, N. (1990). Policy reform, shadow prices, and market prices. *Journal of public economics*, 42(1):1–45.

[40] Duenas, M. and Fagiolo, G. (2014). Global trade imbalances: A network approach. *Advances in Complex Systems*, 17(03n04):1450014.

[41] Erdös, P.; Rényi, A. (1959). On random graphs. *Publ Math*, 6:290–297.

[42] European Union (2015). The euro. Online: europa.eu/about-eu/basic-information/money/euro/index_en.htm. Accessed: 2015-04-21.

[43] Fagiolo, G., Reyes, J., and Schiavo, S. (2008). On the topological properties of the world trade web: A weighted network analysis. *Physica A*, 387(15):3868–3873.

[44] Fagiolo, G., Reyes, J., and Schiavo, S. (2009). World-trade web: Topological properties, dynamics, and evolution. *Physical Review E*, 79(3):036115.

[45] Fagiolo, G., Reyes, J., and Schiavo, S. (2010). The evolution of the world trade web: A weighted-network analysis. *Journal of Evolutionary Economics*, 20(4):479–514.

[46] FAO (2008). Gviews, crop prospects and food situation 3. Food and Agriculture Organization of the United Nations. Online: fao.org/GIEWS/ENGLISH/cpfs/index.htm. Accessed: 2015-09-12.

[47] FAO (2008). The state of food insecurity in the world: High food prices and food security, threats and opportunities. Online: fao.org/docrep/011/i0291e/i0291e00.htm. Accessed: 2015-01-23.

[48] FAO (2010). Price Volatility in Agricultural Markets: Evidence, Impact on Food Security and Policy Responses. Technical report, Food and Agricultural Organization of the United Nations, Rome, Italy.

[49] FAO (2015). FAOSTAT database. Online: faostat.fao.org/. Accessed: 2015-02-23.

www.manaraa.com

[50] Fasano, G. and Franceschini, A. (1987). A multidimensional version of the kolmogorov–smirnov test. *Monthly Notices of the Royal Astronomical Society*, 225(1):155–170.

[51] Fernandez, M., Galeano, J., and Hidalgo, C. (2011). A complex network approach to international commodity trade markets. *International Journal of Complex Systems in Science*, 1(2):191–201.

[52] Fernandez, M., Galeano, J., and Hidalgo, C. (2012). Bipartite networks provide new insights on international trade markets. *Networks and Heterogeneous Media*, 7(3):399–413.

[53] Fisher, I. (1926). *Mathematical investigations in the theory of value and prices*. PhD thesis, Yale University, CT.

[54] Foti, N. J., Pauls, S., and Rockmore, D. N. (2013). Stability of the world trade web over time–an extinction analysis. *Journal of Economic Dynamics and Control*, 37(9):1889–1910.

[55] Fox Keller, E. (2005). Revisiting "scale-free" networks. *BioEssays*, 27(10):1060–1068.

[56] Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239.

[57] Freeman, L. C. (1980). The gatekeeper, pair-dependency and structural centrality. *Quality and Quantity*, 14(4):585–592.

[58] Fu, C. and Wang, X. (2011). Network growth under the constraint of synchronization stability. *Physical Review E*, 83(6):066101.

[59] Gale, D. (1976). The linear exchange model. *Journal of Mathematical Economics*, 3(2):205 – 209.

[60] Garlaschelli, D., Di Matteo, T., Aste, T., Caldarelli, G., and Loffredo, M. I. (2007). Interplay between topology and dynamics in the world trade web. *The European Physical Journal B*, 57(2):159–164.

[61] Garlaschelli, D. and Loffredo, M. I. (2004). Fitness-dependent topological properties of the world trade web. *Physical review letters*, 93(18):188701.

[62] Garlaschelli, D. and Loffredo, M. I. (2005). Structure and evolution of the world trade network. *Physica A*, 355(1):138–144.

[63] Gilarranz, L. L. J., Pastor, J. M. J., and Galeano, J. (2012). The architecture of weighted mutualistic networks. *Oikos*, 121(7):1154–1162.

[64] Gilbert, C. L. (2010). How to understand high food prices. *Journal of Agricultural Economics*, 61(2):398–425.

[65] Gilbert, C. L. and Morgan, C. W. (2010). Food price volatility. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1554):3023–3034.

[66] Gilbert, E. N. (1959). Random graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144.

www.manaraa.com

[67] Giordani, P., Rocha, N., and Ruta, M. (2012). Food prices and the multiplier effect of export policy.

[68] Glick, R. and Rose, A. K. (1999). Contagion and trade: Why are currency crises regional? *Journal of International Money and Finance*, 18(4):603–617.

[69] Goldberg, A. V. (1985). *A new max-flow algorithm*. Laboratory for Computer Science, Massachusetts Institute of Technology, MA.

[70] Good, D. and Irwin, S. (2015). The relationship between stocks-to-use and soybean prices revisited. *Farmdoc Daily*, 5(5: 89).

[71] Götz, L. and Goychuk, K. (2013). Export Restrictions and Market Uncertainty: Evidence from the Analysis of Price Volatility in the Ukrainian Wheat Market. *2013 Annual Meeting, . . .*, pages 1–21.

[72] Guillaume, J.-L. and Latapy, M. (2004). Bipartite structure of all complex networks. *Information processing letters*, 90(5):215–221.

[73] Hazell, P. B. (2009). *The Asian green revolution*, volume 911. Intl Food Policy Res Inst.

[74] Headey, D. and Fan, S. (2008). Anatomy of a crisis: the causes and consequences of surging food prices. *Agricultural Economics*, 39:375–391.

[75] Headey, D. and Fan, S. (2010). *Reflections on the Global Food Crisis: How Did It Happen? How Has It Hurt? And How Can We Prevent the Next One?*, volume 372 of *IFPRI research monograph*. International Food Policy Research Institute.

[76] Henriet, F., Hallegatte, S., and Tabourier, L. (2012). Firm-network characteristics and economic robustness to natural disasters. *Journal of Economic Dynamics and Control*, 36(1):150–167.

[77] Hepburn, A. (2012). What is to blame for rising food prices? Macleans. Online: macleans.ca/economy/business/food-more-expensive-because-supplies-are-low/. Accessed: 2015-01-13.

[78] Ianchovichina, E. and Loening, J. C. (2012). Wood, How vulnerable are Arab countries to global food price shocks? *World Bank Policy Research Working Paper*, (6018).

[79] IMF (2014). International Financial Statistics (1948-2014). *Mimas, University of Manchester*.

[80] International Energy Agency (2015). Energy statistics. Online: www.iea.org/statistics/. Accessed: 2015-10-13.

[81] International Monetary Fund (2001). Global Trade Liberalization and the Developing Countries. Online: imf.org/external/np/exr/ib/2001/110801.htm. Accessed 26-04-16.

[82] International Monetary Fund (2015). International financial statistics database. Online: data.imf.org/. Accessed: 2015-03-10.

[83] Jeffreys, H. (1973). *Scientific inference*. Cambridge University Press, Cambridge, UK.

www.manaraa.com

[84] Kakade, S., Kearns, M., Ortiz, L., Pemantle, R., and Suri, S. (2004a). The economics of social networks. *Proceedings of NIPS*.

[85] Kakade, S. M., Kearns, M., and Ortiz, L. E. (2004b). Graphical economics. In *Learning Theory*, pages 17–32. Springer.

[86] Kakade, S. M., Kearns, M., Ortiz, L. E., Pemantle, R., and Suri, S. (2004c). Economic properties of social networks. In *Advances in Neural Information Processing Systems*, pages 633–640.

[87] Kali, R. and Reyes, J. (2005). Financial Contagion on the International Trade Network. *Business*, 48(4):1072–1101.

[88] Kim, B. J., Trusina, A., Minnhagen, P., and Sneppen, K. (2005). Self organized scale-free networks from merging and regeneration. *The European Physical Journal B*, 43(3):369–372.

[89] König, M. and Battiston, S. (2009). From graph theory to models of economic networks. a tutorial. In *Networks, Topology and Dynamics*, pages 23–63. Springer.

[90] Lampietti, J., Larson, D. F., Battat, M., Erekat, D., de Hartog, A., and Michaels, S. (2011). The Grain Chain : Food Security and Managing Wheat Imports in Arab Countries. *Smart Lessons*, (12):1–4.

[91] Lewis, M. (2005). *The History & Development of Commodity Exchanges*. An Investor Guide To Commodities. Springer, Berlin, Germany.

[92] Lind, P. G., González, M. C., and Herrmann, H. J. (2005). Cycles and clustering in bipartite networks. *Physical review E*, 72(5):056127.

[93] Lipsky, J. (2008). *Commodity prices and global inflation*. Council on Foreign Relations, New York, NY.

[94] MacKenzie, C. (2014). Summarizing Risk Using Risk Measures and Risk Indices. *Risk Analysis*, 34(12):2143–2162.

[95] Maddison, A. (2007). *The World Economy: A millennial perspective and historical statistics*. Academic Foundation.

[96] Magnan, N., Lybbert, T. J., McCalla, A. F., and Lampietti, J. A. (2011). Modeling the limitations and implicit costs of cereal self-sufficiency: the case of Morocco. *Food Security*, 3(S1):49–60.

[97] Martin, W. and Anderson, K. (2011). Export Restrictions and Price Insulation During Commodity Price Booms. *American Journal of Agricultural Economics*, 94(2):422–427.

[98] MATLAB (2015). *Bioinformatics Toolbox R2015a*. The MathWorks Inc., Natick, MA.

[99] Merriam Webster (2015). Merriam Webster Dictionary. Online: merriam-webster.com/dictionary/. Accessed: 2015-06-11.

[100] Milberg, W. (1996). The rhetoric of policy relevance in international economics. *Journal of Economic Methodology*, 3(2):237–259.

www.manaraa.com

[101] Mitchell, D. (2008). A note on rising food prices. Technical report, World Bank Policy Research Working Paper Series 4682. World Bank, Washington, D.C.

[102] Mitra, S. and Josling, T. (2009). Agricultural export restrictions: Welfare implications and trade disciplines. *International Food & Agricultural Trade Policy Council, Washington, DC*, (January).

[103] Mundial, B. (2008). Double jeopardy: responding to high food and fuel prices. *Cumbre Hokkaido-Toyako del G*, 8:2.

[104] Murphy, B. S. and Policy, T. (2009). Strategic Grain Reserves In an Era of Volatility. *Institute for Agriculture and Trade Policy*, (October).

[105] Newman, M. E. (2008). The mathematics of networks. *The new palgrave encyclopedia of economics*, 2(2008):1–12.

[106] Newman, M. E., Forrest, S., and Balthrop, J. (2002). Email networks and the spread of computer viruses. *Physical Review E*, 66(3):035101.

[107] Newman, M. M. E. J. (2003). The structure and function of complex networks. *SIAM review*, 45(2):167–256.

[108] OECD (2014). *OECD Factbook 2014: Economic, Environmental and Social Statistics*. OECD Publishing.

[109] OECD-FAO (2008). OECD-FAO agricultural outlook 2008-2017. pages 1–73.

[110] Ohlin, B. (1967). Interregional and international trade.

[111] Piesse, J. and Thirtle, C. (2009). Three bubbles and a panic: An explanatory review of recent food commodity price events. *Food Policy*, 34(2):119–129.

[112] Prakash, A. (2011). *Safeguarding food security in volatile global markets*. Food and Agriculture Organization of the United Nations, Rome, Italy.

[113] Programme, W. F. (2010). Yemen - Market Study. Technical Report December 2010, World Food Programme, Rome, Italy.

[114] Redner, S. (2005). Citation statistics from more than a century of physical review. In *APS Meeting Abstracts*, volume 1, page 19005.

[115] Ricardo, D., Gonner, E. C. K., and Li, Q. (1819). *The principles of political economy and taxation*. World Scientific, Singapore.

[116] Rojko, A. S. (1975). The economics of food reserve systems. *American Journal of Agricultural Economics*, 57(5):866–872.

[117] Rosegrant, M. W. (2008). *Biofuels and grain prices: impacts and policy responses*. International Food Policy Research Institute Washington, DC.

[118] Sadler, M. and Magnan, N. (2011). Grain import dependency in the MENA region: risk management options. *Food Security*, 3(S1):77–89.

www.manaraa.com

[119] Schmidhuber, J. (2008). Impact of an increased biomass use on agricultural markets, prices and food security: A longer-term perspective. In *Energy Security in Europe: Proceedings from the Conference 'Energy Security in Europe'*, pages 133–170.

[120] Schnepf, R. (2008). High agricultural commodity prices: What are the issues? Library of Congress, Washington, DC.

[121] Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A., and White, D. R. (2009). Economic networks: The new challenges. *Science*, 325:422–425.

[122] Serrano, M. A. and Boguñá, M. (2003). Topology of the world trade web. *Physical Review E*, 68(1 Pt 2):015101.

[123] Serrano, M. A., Boguñá, M., and Vespignani, A. (2007). Patterns of dominant flows in the world trade web. *Journal of Economic Interaction and Coordination*, 2(2):111–124.

[124] Shama, R. (2011). *Food export restrictions: Review of the 2007-2010 experience and considerations for disciplining restrictive measures*. Number 32. FAO, Rome, Italy.

[125] Sharma, R. (2011). ENHANCING MARKET. Technical report, FAO: Agricultural Market Information System.

[126] Sinha, S., Chatterjee, A., Chakraborti, A., and Chakrabarti, B. (2010). *Econophysics: an introduction*. John Wiley & Sons, Hoboken NJ.

[127] Slayton, T. (2009). Rice crisis forensics: How Asian governments carelessly set the world rice market on fire. *Center for Global Development working paper*, (163).

[128] Squartini, T., Fagiolo, G., and Garlaschelli, D. (2011a). Randomizing world trade. i. a binary network analysis. *Physical Review E*, 84(4):046117.

[129] Squartini, T., Fagiolo, G., and Garlaschelli, D. (2011b). Randomizing world trade. ii. a weighted network analysis. *Physical Review E*, 84(4):046118.

[130] Stumpf, M. P. and Porter, M. A. (2012). Critical truths about power laws. *Science*, 335(6069):665–666.

[131] Subramanian, A. and Wei, S.-J. (2007). The wto promotes trade, strongly but unevenly. *Journal of international Economics*, 72(1):151–175.

[132] Suri, S. (2007). *The effects of network topology on strategic behavior*. PhD thesis, University of Pennsylvania, Philadelphia, PA.

[133] The Federal Reserve Bank of New York (2015). Historical us dollar to euro exchange rate. http://www.newyorkfed.org/data-and-statistics/index.html. Accessed: 2015-09-12.

[134] Timmer, C. (2008). *Causes of high food prices*. Asian Development Bank, Mandaluyong, Philippines.

[135] Tinbergen, J. (1962). *An analysis of world trade flows. Shaping the world economy*. Twentieth Century Fund, New York, NY.

www.manaraa.com

[136] Torero, M. and von Braun, J. (2004). Alternative mechanisms to reduce food price volatility and price spikes. *Chemistry & biodiversity*, 1(11):1829–1841.

[137] Tradingcharts.com (2015). The stocks to use ratio. Online: futures.tradingcharts.com. Accessed: 2015-06-22.

[138] Trostle, R. (2010). *Global Agricultural Supply and Demand: Factors Contributing to the Recent Increase in Food Commodity Prices*. DIANE Publishing, Collingdale, PA.

[139] UN Department of Economic and Social Affairs, Statistics Division (2010). International Merchandise Trade Statistics: Concepts and Definitions 2010. Technical report, The United Nations, ROME.

[140] United Nations (2009). United nations millennium development goals. Online: www.un.org/millenniumgoals/. Accessed: 2015-02-09.

[141] United Nations Statistics Division, New York, NY (2014). United Nations Commodity Trade Statistics Database (UN comtrade). Online: comtrade.un.org/. Accessed: 2015-06-22.

[142] USDA (2011). *USDA Agricultural Projections To 2017*. Long-term Projections Report OCE-2008-1, U.S. Department of Agriculture, Washington, D.C.

[143] USDA, E. R. S. (2015). Yearbook data. Online: ers.usda.gov/data-products/. Accessed: 2015-09-12.

[144] von Braun, J., Torero, M., Braun, J. V., and Torero, M. (2009). Implementing physical and virtual food reserves to protect the poor and prevent market failure. *IFPRI Policy Brief 10*, February(No. 04):1–4.

[145] World Bank (2005a). Household energy supply and use in Yemen. Technical Report Energy Sector Management Assistance Program working paper series; ESM 315/05, World Bank, Washington, DC.

[146] World Bank (2005b). *Managing Food Price Risks and Instability in an Environment of Market Liberalization*. Number 32727. World Bank Washington, DC.

[147] World Bank (2012). The Grain Chain: Food Security and Managing Wheat Imports in Arab Countries. Technical report, World Bank, Washington, DC.

[148] World Bank (2014). World development indicators 2014. Online: data.worldbank.org. Accessed: 2015-06-11.

[149] World Trade Organization (2013). WTO Trade Monitoring database. Accessed: 2015-06-12.

[150] Wright, B. (2009). International Grain Reserves: And Other Instruments to Address Volatility in Grain Markets. (August).

[151] Wright, B. D. (2011). The Economics of Grain Price Volatility. *Applied Economic Perspectives and Policy*, 33(1):32–58.

[152] Wright, B. D. (2012). International grain reserves and other instruments to address volatility in grain markets. *World Bank Research Observer*, 27(2):222–260.

www.manaraa.com

[153] Wu, F. and Guclu, H. (2013). Global maize trade and food security: implications from a social network model. *Risk analysis*, 33(12):2168–78.

[154] Yang, X.-H., Lou, S.-L., Chen, G., Chen, S.-Y., and Huang, W. (2013). Scale-free networks via attaching to random neighbors. *Physica A*, 392(17):3531–3536.

www.manaraa.com

Appendix A

Addendum to Chapter IV

A.1 Comtrade data to Network Function

Listing A.1 Convert comtrade data to network function

```
function [ A, B ] = csv_to_value_matrix(data,index)
% convert UN Comtrade commodity data (.csv format) to adjacency matrix

% INPUT:
% data - UN comtrade commodity data: import raw .csv as a table <-- important!
% index - list of UN country codes (i.e., 156 = China...)

% OUTPUT:
% A - weighted adjacency matrix of trade (measured in US$)
% B - adjacency matrix of trade

min_year = min(data.Year); % find start year
max_year = max(data.Year); % find end year
A = zeros(length(index), length(index), numel(unique(data.Year)));
B = zeros(length(index), length(index), numel(unique(data.Year)));

for i = 1:numel(unique(data.Year))
y = data.Year==(min_year-1)+i;                          % filter data by year
% Use the ismember function to get the mapping between vertex ids and
% their consecutive index mappings:
[~,from] = ismember(data.ReporterCode(y), index);       % exporter index
[~,to] = ismember(data.PartnerCode(y), index);          % impoerter index
Adj = zeros(length(index), length(index));
% Use sub2ind to populate adjacency matrix:
linear_ind = sub2ind(size(Adj), from, to);
Adj(linear_ind) = data.Value(y);                        % add value of trade
A(:,:,i) = Adj;
```

www.manaraa.com

```
B(:,:,i) = Adj > 0;
end
end
```

A.2 Trade network growth algorithm

Listing A.2 Trade network growth algorithm

```
%% Trade network growth algorithm
% model parameters
time = 1000;                % number of time periods
lambda = 0.1;               % vertex attachment probability
ini_pref = 0.1;             % initial preference, must be > 0

% inistialise graph
u = 1;                      % number of nodes
G = zeros(time,time);       % create seed network
sumedges = sum(sum(G));     % starting edges
G(1,1,:) = 1;

% start time
t = 0;
while t < time
% add 1 to timer
t = t + 1;
%_____EXTRINSIC GROWTH_____%
if rand <= lambda
% add node to U
u = u + 1;
% select node u in U
node_u = u;
for i = 1:5
% attach edge via pref attachment
node_y = pref_attach( G, node_u, ini_pref);
% add edge from node_v to node_u
G(node_y,node_u) = 1;
% add 1 to edge count
sumedges = sumedges + 1;
end
else
%_____INTRINSIC GROWTH_____%
% attach edge via uniform attachment
node_a = ceil(rand * u);
% attach edge via pref attachment
node_b = pref_attach( G, node_a, ini_pref);
% add edge from node_v to node_u
G(node_b,node_a) = 1;
```

www.manaraa.com

```
% add 1 to edge count
sumedges = sumedges + 1;
end
end
```

Listing A.3 Preferential attachment function

```
function [ node_v ] = pref_attach( G, node_u, ini_pref)

% INPUT:
% G – adjacency matrix
% node_n – source node
% ini_pref – initial preference, must be > 0 (hidden model parameter)

% OUTPUT:
% node_m – target node

% find k in G
degree = sum(G(1:node_u,1:node_u),2);
if sum(degree) < size(degree,1)
% check if connected
check_link = false;
while ~check_link
% pick node_v via pref attachment
node = RouletteWheelSelection(degree+ini_pref);
% check if connected
check_link = (G(node,node_u) == 0);
end
node_v = node;
else
% if all are connected
node_v = RouletteWheelSelection(degree+ini_pref);
end
```

www.manaraa.com

A.3 Empirical degree distributions for selected markets

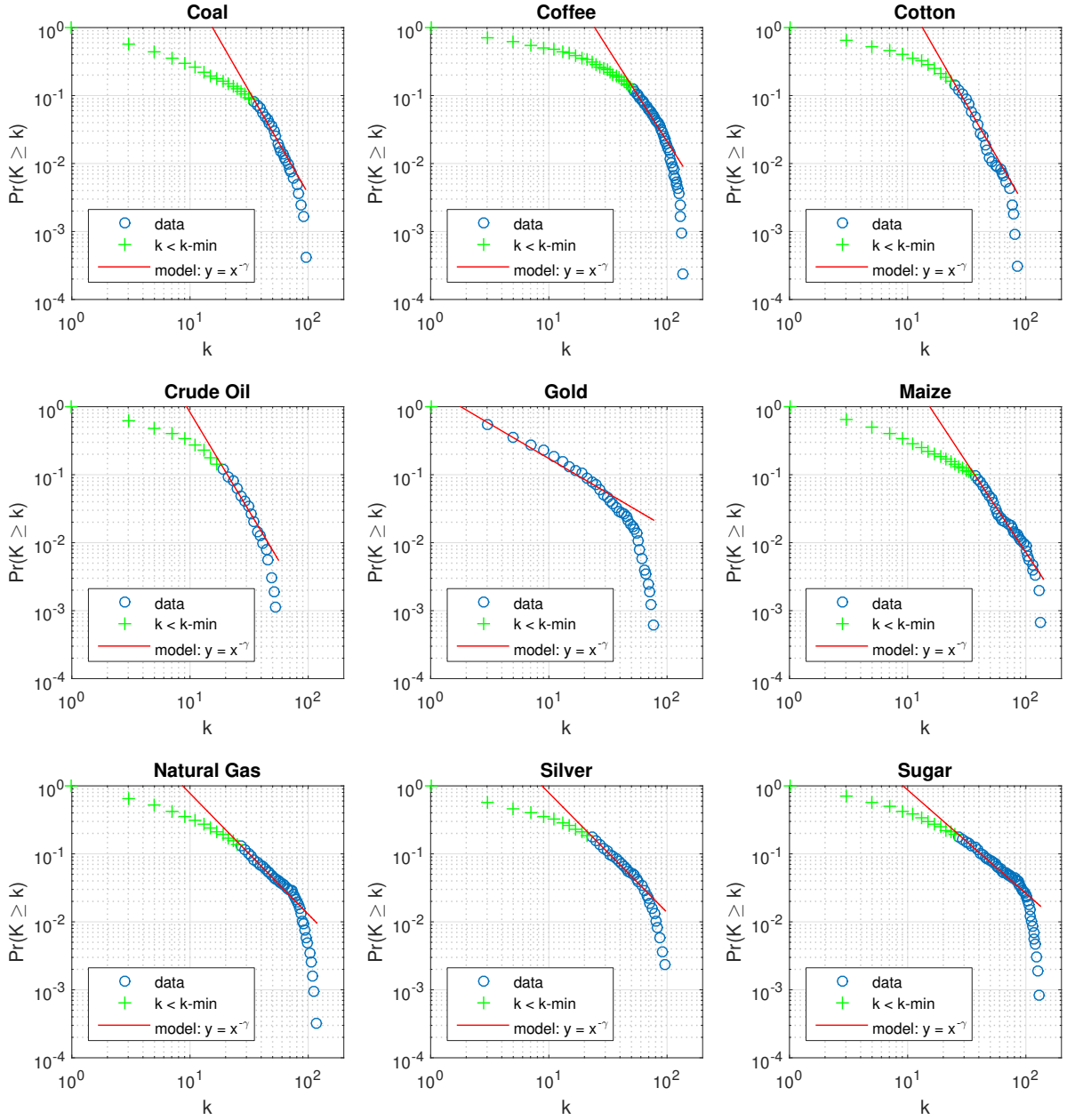


Fig. A.1 Empirical degree distributions (exports) for selected markets. Power-law fit on *log-log* axis. Data source: [141].

www.manaraa.com

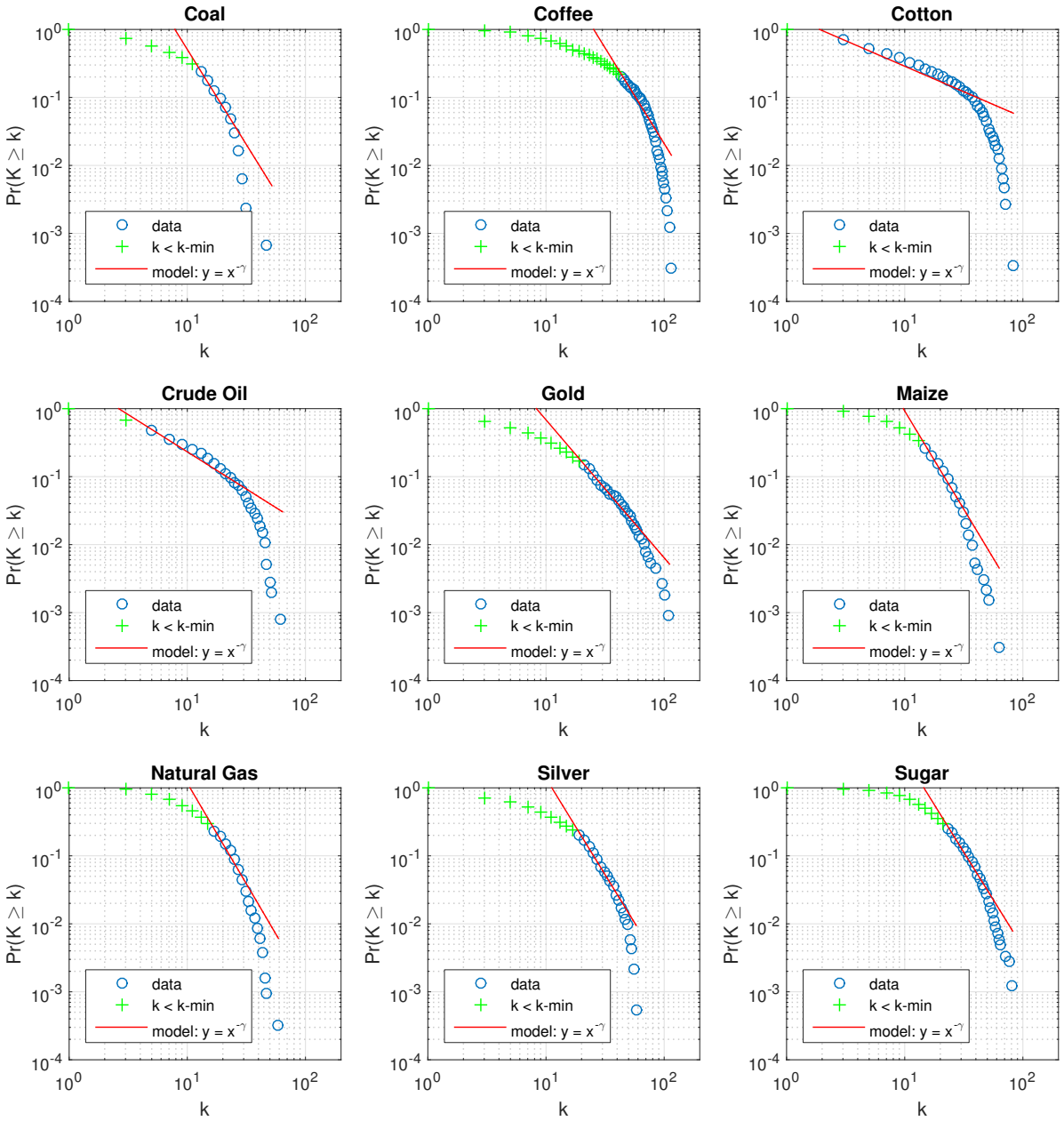


Fig. A.2 Empirical degree distributions (imports) for selected markets. Power-law fit on *log-log* axis. Data source: [141].

www.manaraa.com

Appendix B

Addendum to Chapter V

B.1 The Graphical Exchange Model

Listing B.1 Graphical Exchange Model

```
%% Graphical Exchange Model

% Create random graph
G = rand(10,10);
G(G < 0.5) = 1;
G(G < 1) = 0;

p(1:size(G,1)) = 1/max(sum(G,1));     % price(i) = 1 / max-degree(i)

fill = zeros(size(G,1),size(G,1));  % create empty fill matrix

G = [fill G;                        % make G into bipartite network
fill fill;];

n = size(G,1)/2;                    % number of (n) sellers
m = n;                              % equal number of buyers and sellers

% Add sink and source nodes

[F,s,t] = sink_and_source(G,n,m);

% add source and sink nodes to graph G, such that F =

%%%%%%%%%%%%
% [G        %
%    s      %
%      t];  %
%%%%%%%%%%%%

% s = source node
```

www.manaraa.com

```
% t = sink node

% Set initial prices

% prices (p) are first set to 1/max(k) for each (n) seller

p = round(p/0.01)*0.01;
F(s,1:n) = F(s,1:n) .* p;       % add prices to source nodes

%% Primal-Dual Algorithm

% Primal: For a given set of prices (p), calc max-flow acorss the graph

% Dual: Increase prices untill market reaches equilibrium

%tic; % timer
%tight_prices = primal_dual( F, p );
%toc


t = size(F,1);                          % t = sink node
s = (size(F,1))-1;                      % s = source node
n = (size(F,1)-2)/2;                    % number of (n) sellers
tight_prices = zeros(1,n);              % pre-allocate tight_prices
%price_index = 1:n;                      % index for keeping track of
    prices
Matrix = sparse(F);
%tight = min(tight_prices) > 0;

while sum(Matrix(1)) == 0

tic; % timer
warning('off','bioinfo:graphmaxflow:mincutsLimit')
[~, ~, Cut] = graphmaxflow(Matrix, s, t);
toc

min_cut = find(Cut(1,:) > 0);           % find min-cut
min_cut = min_cut(min_cut~=s);          % remove source node
min_cut = min_cut(min_cut<=n);          % remove source node
cut_size = size(min_cut,2);

Matrix(min_cut,:) = [];     % seller links
Matrix(min_cut+(n-cut_size),:) = [];     % seller links
Matrix(:,min_cut) = [];      % buyer links
Matrix(:,min_cut+(n-cut_size)) = [];      % buyer links
tight_prices(min_cut) = max(p);
p(min_cut) = [];
p = p + 0.01; % increase prices of active sellers
```

www.manaraa.com

```
p = round(p*100000)/100000;
s = (size(Matrix,1))-1;
Matrix(s,1:size(p,2)) = p; % update flow matrix
n = (size(Matrix,1)-2)/2;                    % number of (n) sellers
t = size(Matrix,1);
s = (size(Matrix,1))-1;
%tight = min(tight_prices) > 0;

end

disp(tight_prices) % print prices at equilibrium

%% Re-create flow matrix at equilibrium

t = size(F,1);                          % t = sink node
s = (size(F,1))-1;                      % s = source node
n = (size(F,1)-2)/2;                    % number of (n) sellers
F(s,1:n) = tight_prices;

[~, FlowMatrix] = graphmaxflow(sparse(F), s, t);

F_flow = full(FlowMatrix);
F_flow = round(F_flow*100000)/100000; % 1 does not equal 1.000

view(biograph(F,[],'ShowWeights','on')) % graph visualisation (optional)
```

Listing B.2 Add sink and source function

```
function [ F, s, t ] = sink_and_source( A, n, m )

% modified from: http://www.mathworks.com/matlabcentral/fileexchange
   /46227-seam-carvin
% g-with-dijkstra-and-dynamic-programming/content/AddSrcAndSink2Graph.m

start_nodes = 1:n;              % first col of sellers
end_nodes   = n+1:n+m;          % last col of buyers

[ii,jj,s] = find(A);            % disect current graph
[mG,nG] = size(A);              % find hight and width of A

%adding source node
ii_src = (n+m+1)*ones(m, 1);    % s -> first col
jj_src = start_nodes';
s_src  = ones(m,1);             %equal oppurtunity for entire row

%adding sink node
ii_trg = end_nodes';
jj_trg = (n+m+2)*ones(m, 1);    % last col -> t
s_trg  = ones(m,1);             %equal oppurtunity for entire row
```

www.manaraa.com

```
21
22 %resemble Graph with source and sink
23 F = sparse([ii;ii_src;ii_trg], ...
24 [jj;jj_src;jj_trg], [s;s_src;s_trg], mG+2, nG+2);
25
26 F = full(F);                    % creates the flow graph F
27
28 s = size(F,1)-1;                % s = source node
29 t = size(F,1);                  % t = sink node
30
31 end
```

Listing B.3 Find tight set function

```
1
2 function [ tight_index ] = find_tight( F, n , size_G, s, t)
3
4 % find tight sets in flow matrix -
5
6 % INPUT:
7 % F_flow - adjacency matrix of (n*m) graph G
8 % n - (1*n) price vector
9 % size_G = size of graph
10
11 %% For sellers
12 money_sellers(1:n) = 0;
13
14 for i = 1:n
15 money_sellers(i) = F(s,i);
16 end
17
18 money_sellers = round(money_sellers*100000)/100000; % 1 does not equal
      1.000
19
20 % money_N(j) = sum(e_i) for all (i) in N(j) = total endowment in N(j)
21
22 money_sellers_N(1:n) = 0;
23
24 for i = 1:n
25 N = find(F(i,1:size_G) > 0);
26 money_sellers_N(i) = sum(F(N,t));
27 end
28
29 %% For buyers
30 money_buyers(1:n) = 0;
31
32 for i = 1:n
33 money_buyers(i) = F(i+n,t);
34 end
```

www.manaraa.com

```
35
36 money_buyers = round(money_buyers*100000)/100000; % 1 does not equal
       1.000
37
38 % money_N(j) = sum(e_i) for all (i) in N(j) = total endowment in N(j)
39 money_buyers_N(1:n) = 0;
40
41 for i = 1:n
42 N = find(F(1:size_G,i+n) > 0);
43 money_buyers_N(i) = sum(F(s,N));
44 end
45
46 %% sets
47 tight(1:n+n) = 0; % denotes where j and N(j) is tight
48
49 for i = 1:n
50 if money_sellers(i) == money_sellers_N(i) || money_sellers(i) == 0 &&
       money_sellers_N(i) == 0
51 tight(i) = 1;
52 tight(find(F(i,1:size_G) > 0)) = 1;
53 end
54 end
55
56 for i = 1:n
57 if money_buyers(i) == money_buyers_N(i) || money_buyers(i) == 0 &&
       money_buyers_N(i) == 0
58 tight(i+n) = 1;
59 tight(find(F(1:size_G,i+n) > 0)) = 1;
60 end
61 end
62
63 tight_index = find(tight > 0); % find index of tight sellers
64
65 end
```

www.manaraa.com

Appendix C

Addendum to Chapter VI

C.1 List of Export Restrictions in the Maize Market (2006–2012)

Table C.1 Export Restrictions in the Maize Market (2006–2012)

Country	Start Date	End Date	Measure Type	Reference
Ethiopia	01/01/2007	01/07/2010	Export Ban	[28]
Argentina	01/11/2007		Export Tax	[38]
Argentina	01/12/2007	01/05/2008	Export Restriction	[38]
China	01/01/2008	01/11/2008	Export Tax / Quota	[125]
Ethiopia	01/02/2008		Export Ban	[125]
Tanzania	01/02/2008		Export Ban	[125]
Niger	01/03/2008		Export Ban	[125]
Argentina	01/04/2008		Export Tax	[125]
Bolivia	01/04/2008		Export Ban	[125]
Argentina	01/06/2008	01/09/2008	Export Quota	[125]
China	01/11/2008		Export Tax / Quota	[125]
Bolivia	24/02/2010	06/05/2010	Export Restriction	[149]
Russian Federation	15/07/2010	31/12/2010	Export Ban	[28]
Ukraine	04/10/2010	05/05/2011	Export Restriction	[28]
Russian Federation	02/01/2011	30/06/2011	Export Ban	[28]
Belarus	12/06/2011		Export Ban	[28]

www.manaraa.com

C.2 List of Export Restrictions in the Wheat Market (2006–2012)

Table C.2 Export Restrictions in the Wheat Market (2006–2012)

Country	Start Date	End Date	Measure Type	Reference
Ukraine	01/10/2006	01/04/2007	Export Restriction	[38]
Ethiopia	01/01/2007	01/07/2010	Export Ban	[28]
Argentina	01/03/2007	01/11/2007	Export Restriction	[38]
Ukraine	01/06/2007	01/06/2008	Export Ban	[125]
India	01/10/2007	09/09/2011	Export Restriction	[28]
Argentina	01/12/2007	01/05/2008	Export Restriction	[38]
China	01/01/2008	01/11/2008	Export Tax / Quota	[125]
Ethiopia	01/02/2008		Export Ban	[125]
Russian Federation	01/02/2008		Export Tax	[125]
Tanzania	01/02/2008		Export Ban	[125]
Kazakhstan	01/03/2008		Export Tax	[38]
Niger	01/03/2008		Export Ban	[38]
Pakistan	31/03/2008	01/10/2009	Export Tax	[125]
India	03/07/2009	31/03/2013	Export Restriction	[149]
Pakistan	08/02/2010		Export Restriction	[149]
Argentina	01/05/2010	01/01/2011	Export Quota	[125]
Pakistan	30/06/2010	13/12/2010	Export Ban	[28]
Russian Federation	15/07/2010	31/12/2010	Export Ban	[28]
Ukraine	04/10/2010	05/05/2011	Export Restriction	[28]
Ukraine	28/12/2010	30/06/2011	Export Restriction	[149]
Russian Federation	02/01/2011	30/06/2011	Export Ban	[28]
Moldova	03/02/2011	04/05/2011	Export Ban	[28]
Serbia	16/03/2011	16/06/2011	Export Restriction	[149]
Macedonia	08/04/2011	15/09/2011	Export Restriction	[149]
Belarus	12/06/2011		Export Ban	[28]
India	09/09/2011		Export Restriction	[149]
Ukraine	19/10/2011	04/10/2011	Export Restriction	[149]
India	23/02/2012		Export Restriction	[149]
Iran	01/11/2012		Export Restriction	[149]

www.manaraa.com

C.3 List of Export Restrictions in the Rice Market (2006–2012)

Table C.3 Export Restrictions in the Rice Market (2006–2012)

Country	Start Date	End Date	Measure Type	Reference
Argentina	01/04/2008		Export Ban	[125]
Bangladesh	01/05/2008	07/11/2008	Export Ban	[125]
Belarus	12/06/2011		Export Ban	[28]
Bolivia	01/04/2008		Export Ban	[125]
Bolivia	14/03/2012		Export Restriction	[149]
Cambodia	01/04/2008	01/06/2008	Export Ban	[125]
China	01/01/2008	01/11/2008	Export Tax / Quota	[125]
Egypt	19/01/2008	12/02/2008	Export Ban	[127]
Egypt	01/04/2008	01/10/2012	Export Ban	[127]
Ethiopia	01/01/2007	01/07/2010	Export Ban	[28]
India	31/10/2007		Export Tax	[125]
India	01/12/2007		Export Tax	[125]
India	27/03/2008		Export Tax	[125]
India	03/06/2010	09/09/2011	Export Restriction	[149]
India	21/02/2012	04/07/2012	Export Restriction	[149]
India	10/05/2012		Export Restriction	[149]
Indonesia	11/04/2008		Export Restriction	[125]
Madagascar	01/05/2008		Export Ban	[125]
Nepal	31/03/2008	01/11/2008	Export Ban	[125]
Niger	01/03/2008		Export Ban	[125]
Pakistan	30/04/2008	01/10/2008	Export Restriction	[127]
Tanzania	01/02/2008		Export Ban	[125]
Thailand	01/05/2008		Export Ban	[125]
Vietnam	21/07/2007		Export Ban	[127]
Vietnam	18/01/2008		Export Tax	[127]
Vietnam	05/02/2008	18/06/2008	Export Ban	[127]
Vietnam	18/06/2008		Export Restriction	[127]

www.manaraa.com

C.4 Convert Comtrade Data to Weighted Network Function

Listing C.1 Convert comtrade data to weighted network function

```
function [ A, B ] = csv_to_weight_matrix(data,index)
% convert UN Comtrade commodity data (.csv format) to matrix

% INPUT:
% data – UN comtrade data: import .csv as a table <-- important!
% index – list of UN country codes selected (i.e., 156 = China...)

% OUTPUT:
% A – weighted adjacency matrix of trade (measured in kg)
% B – adjacency matrix of trade

min_year = min(data.Year); % find start year
max_year = max(data.Year); % find end year
A = zeros(length(index), length(index), numel(unique(data.Year)));
B = zeros(length(index), length(index), numel(unique(data.Year)));

for i = 1:numel(unique(data.Year))
y = data.Year==(min_year-1)+i;                          % filter data by
   year
% Use the ismember function to get the mapping between vertex ids and
% their consecutive index mappings:
[~,from] = ismember(data.ReporterCode(y), index);       % exporter index
[~,to] = ismember(data.PartnerCode(y), index);          % impoerter index
Adj = zeros(length(index), length(index));
% Use sub2ind to populate adjacency matrix:
linear_ind = sub2ind(size(Adj), from, to);
data.Netweightkg(isnan(data.Netweightkg)) = 0 ; % clear NaNs
Adj(linear_ind) = data.Netweightkg(y);                          % add value
    of trade
A(:,:,i) = Adj;
B(:,:,i) = Adj > 0;
end

end
```

www.manaraa.com

Appendix D

Scientific Contributions

D.1 List of invited presentations

Royal Academy of Engineering. 3rd Mar 2016. "*Complexity Science Tools for Environmental Policy and Practice.*" Demonstrating Complexity Science Tool-kits to government policy-makers. Prince Philip House, London

University College London. 9th-10th Nov 2015. "*Price Variation in Network Models of International Food-Commodity Markets.*" BHP Billiton Sustainable Communities/UCL Grand Challenges Symposium on Global Food Security: Adaptation, Resilience and Risk. UCL, London

Cornell University. 11th-14th Oct 2015. "*The Network Topology of Risk.*" 2nd International Conference on Global Food Security. Ithaca, New York, USA

University of Warwick. 15th-17th Apr 2015. "*A Network Data Analysis of Global Food-Commodity Markets.*" Bayesian Decision Support for Food Security. University of Warwick, Department of Statistics

University College London. 23rd May 2014. "*Failure Cascades in Real-World Networks.*" Complexity Science in the Real World (CSRW): PhD Conference & Hackathon. UCL, London

www.manaraa.com

Royal Society. 25th Apr 2014. "*Cascading Failure in Networked Systems.*" Complexity Science in the Real World (CSRW): Models for Real World Policy. Royal Society, London

Chicheley Hall. 3rd July 2013. "*Modelling Path Dependency and Inoperability in Global Supply Chains.*" Complexity Science in the Real World (CSRW): Modelling Workshop. Chicheley Hall, Buckinghamshire

www.manaraa.com