



Universität Potsdam  
Mathematisch-Naturwissenschaftliche Fakultät



## Master's Thesis

# Optimal adjustment of the global trade system to local network disruption

# Optimale Anpassung des Welthandelssystems an lokale Netzwerkstörungen

prepared by

**Sebastian Klipp**

at the University of Potsdam

and the Potsdam Institute for Climate Impact Research

**Thesis period:** October 2015 until April 2016

**Supervisor:** M.Sc. Leonie Wenz

**First Referee:** Prof. Dr. Anders Levermann

**Second Referee:** Jun.-Prof. Dr. Nicole Glanemann



## **Abstract**

Trade networks represent monetary flows between different regions and sectors globally. If production in a specific regional sector is reduced, e.g. due to extreme weather, the network has to adapt in order to compensate for the failure. In this thesis, an optimal network response is investigated, minimising the sum of deviations from the initial production of each regional sector and satisfying the constant final demand of all regions. A framework for this linear optimisation problem is provided, finding that the response of the network is linear up to a threshold that is specific to the failing regional sector. In most cases there is one primary compensator that solely increases its production significantly. Three main node characteristics are identified that serve as predictors for the primary compensatory node for a given failure. The most important characteristic is the strength of the regional output, whereas the number of shared output neighbours with the failing node and the regional sector's input-output ratio play a crucial role for specific sectors only.

## **Zusammenfassung**

Handelsnetzwerke repräsentieren monetäre Flüsse zwischen einzelnen Regionen und wirtschaftlichen Sektoren weltweit. Wird die Produktion eines Regionalsektors reduziert, z.B. in Folge von extremen Wetterereignissen, so muss sich das Netzwerk an diesen Ausfall anpassen. In dieser Arbeit wird eine optimale Antwort des Netzwerkes untersucht, wobei die Summe der Abweichungen von der ursprünglichen Produktion eines jeden Regionalsektors minimiert und gleichzeitig die konstante Konsumnachfrage jeder Region erfüllt wird. Die Reaktion des Netzwerkes ist bis zu einem vom ausfallenden Regionalsektor abhängigen Grenzwert linear und in den meisten Fällen gibt es einen Hauptkompensator, der seine Produktion als einziger signifikant erhöht. Es werden drei Eigenschaften von Netzwerknoten bestimmt, die als Prädiktoren für den Hauptkompensator wirken. Die wichtigste Eigenschaft ist die Gesamtproduktionsmenge eines Regionalsektors, während in einzelnen Sektoren die Anzahl der gemeinsamen Nachbarn der ausgehenden Flüsse und das Input-Output-Verhältnis eine wichtige Rolle spielen.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Introduction . . . . .	1
<b>2</b>	<b>Data, methods and theory</b>	<b>3</b>
2.1	Data: Multi-regional input-output tables (MRIOTs) . . . . .	3
2.1.1	Introduction to MRIOTs . . . . .	3
2.1.2	The EORA world MRIO database . . . . .	4
2.2	Methods: Linear Optimisation . . . . .	5
2.2.1	Linear optimisation problems . . . . .	5
2.2.2	Simplex algorithm . . . . .	7
2.2.3	Absolute value target function . . . . .	13
2.3	Theory: Network properties . . . . .	14
<b>3</b>	<b>Problem description</b>	<b>17</b>
3.1	Optimal trade network response to production failure . . . . .	17
<b>4</b>	<b>Model setup</b>	<b>21</b>
4.1	Formulation of the linear optimisation problem . . . . .	21
4.1.1	Free variables and target function . . . . .	21
4.1.2	General linear problem (LPG) . . . . .	22
4.1.3	Special linear problem (LPS) . . . . .	25
4.2	Preprocessing of EORA data . . . . .	27
4.3	Aggregated version of EORA data . . . . .	29
4.4	Technical implementation . . . . .	30
<b>5</b>	<b>Application of the linear problem to a stylised network</b>	<b>31</b>
5.1	Exemplary optimisation results . . . . .	31
<b>6</b>	<b>Optimisation results</b>	<b>35</b>
6.1	Adaptation effort for different failure strengths . . . . .	35
6.2	Adaptation effort for different failure locations . . . . .	40
6.3	Main compensators . . . . .	43
<b>7</b>	<b>Network properties as main compensator predictors</b>	<b>47</b>
7.1	Node property analysis . . . . .	47

*Contents*

7.2 Main compensator probability calculation . . . . .	50
7.3 Main compensator predictors . . . . .	52
<b>8 Conclusions</b>	<b>57</b>
8.1 Summary and Discussion . . . . .	57
8.2 Outlook . . . . .	59
<b>Appendix</b>	<b>61</b>
A.1 Sectors and regions . . . . .	61
A.2 Network adaptation effort $F(i^*r^*)$ . . . . .	63
A.3 Main compensator overviews . . . . .	65
A.4 Main compensator probabilities . . . . .	68

# Nomenclature

## Model variables

Variable	Meaning	Unit
$n_r$	Number of regions	Quantity
$n_s$	Number of economic sectors	Quantity
$n_{rs}$	Number of regional sectors	Quantity
$n_Z$	Number of intermediate flows	Quantity
$n_Y$	Number of final demand flows	Quantity
$\epsilon$	Exogenous forcing strength	Quantity
$X_{ir}$	Total production of regional sector $ir$	$10^3 \$/\text{year}$
$p_{ir}$	Production ratio of regional sector $ir$	Quantity
$\alpha_{ir}$	Efficiency of regional sector $ir$	Quantity
$Z_{ir \rightarrow js}$	Intermediate flow from $ir$ to $js$	$10^3 \$/\text{year}$
$q_{ir \rightarrow js}$	Intermediate flow ratio of flow from $ir$ to $js$	Quantity
$Y_{ir \rightarrow s}$	Final demand flow from $ir$ to $s$	$10^3 \$/\text{year}$
$q_{ir \rightarrow s}$	Final demand flow ratio of flow from $ir$ to $s$	Quantity
$N_{ir,js}^{out}$	Shared output neighbours between $ir$ and $js$	Quantity

## Indices

Index	Meaning
$i, j$	Economic sectors $i, j \in \{1, \dots, n_s\}$
$r, s$	Regions $r, s \in \{1, \dots, n_r\}$
$ir, js$	Regional sectors of sector $i, j$ in region $r, s$
$i^*r^*$	Regional sector affected by exogenous forcing $\epsilon$



# 1 Introduction

## 1.1 Introduction

Extreme events like droughts, floods or storms are likely to increase in intensity and frequency as the global mean surface temperature rises with ongoing climate change (Rahmstorf and Coumou, 2011). The *Intergovernmental Panel on Climate Change (IPCC)* emphasises the relevance of extreme events in a special report (IPCC, 2012), as they may cause destruction and humanitarian emergencies locally in the affected region. In this context, climate change possibly contributes to economic losses (Mechler and Bouwer, 2015) since disasters may cause economic damages in local industries leading to a temporary reduction of production and transport (van der Veen, 2004). In 2013 a once-in-a-hundred-years flood impacted Germany, causing damages estimated as \$6.2 billion (Schulte in den Bäumen et al., 2015), and in the US droughts and heat waves have diminished crop yields (Troy et al., 2015). Parallel to an increase of extreme events the world economy has become increasingly interconnected with global supply chains spanning multiple countries. In a more globalised world, disruptions do not necessarily stay local, but can propagate along supply chains through the trade network (Levermann, 2014; Escaith, 2009). Highly connected macroeconomic systems also play an important role in risk assessments, e.g. in studies focussing on interdependent infrastructures (Haimes et al., 2008).

One method to analyse trade flows in the context of disaster effects is the Input-Output-Analysis (IOA) (Leontief, 1986), which has been used for instance to analyse the outcomes of the 2011 Japanese earthquake and tsunami on international production (MacKenzie et al., 2012). A similar technique was applied to conduct a risk assessment of Hurricane Katrina (Crowther et al., 2007), which hit the US coast in 2005. The economic effects of Katrina were also studied through an adaptive regional input-output model (Hallegatte, 2008) with the outcome that total losses increased nonlinearly with respect to direct losses and that also backward propagation effects play a role for disaster assessments. A comprehensive study of the estimated

## 1 Introduction

economic impacts of natural disasters was conducted by assuming an inter-sectoral model combined with an IO analysis (Okuyama et al., 2004).

The model *Acclimate* is used for studies of dynamic economic loss propagation through the global trade system (Bierkandt et al., 2014; Wenz et al., 2014). In this multi-agent model, each country's industries find an optimal response to a local network disruption, considering revenue maximisation and transportation cost and using multi-regional input-output tables (MRIOTs) (Otto et al., 2016). While classical Input-Output-Analysis only provides static economic loss estimations, *Acclimate* allows to investigate dynamic loss propagation and recovery times.

The accuracy of a global trade flow analysis depends strongly on the regional and sectoral resolution of the available data. The project *zeean* ([zeean.net](http://zeean.net)) aims at providing an MRIOT data set at high regional, sectoral and temporal detail, representing 3051 regions and 419 industries.

Computable General Equilibrium (CGE) models also make use of IO data, trying to estimate a potential response of the economy to external factors by solving systems of economic equations, e.g. in the form of an optimisation problem (Nordhaus and Yang, 1996). Despite their broad utilisation in various fields (Berrittella et al., 2006; Eboli et al., 2010; Nagurney et al., 2002), few studies connect optimisation-based CGE models with global MRIO data (as e.g. the GTAP database and model does (Narayanan et al., 2015; Hertel, 1997)).

The contribution of this thesis to the topic is the provision of a general framework to study an optimal response of the global trade system to local production defects for the case that the deviation from the initial production is minimal and that final consumption is still satisfied. Multi-regional input-output tables are combined with a linear optimisation approach to determine the optimal network response, which is thereupon investigated with regard to particular network properties in order to understand the response behaviour. These studies contribute to a better comprehension of the global trade network's adaptive capacity and help identifying critical regional industries that are crucial for stable global trade flows.

Chapter 2 provides theoretical and methodological backgrounds, whereas the description, formulation and an exemplary application of the optimisation problem are covered by chapters 3, 4 and 5. The analysis of the optimal responses is given in chapter 6, while a technique to estimate the influence of network properties on the optimisation results is developed and applied in chapter 7. Chapter 8 concludes with a discussion and an outlook.

## 2 Data, methods and theory

### 2.1 Data: Multi-regional input-output tables (MRIOTs)

#### 2.1.1 Introduction to MRIOTs

The description of macroeconomic interrelations between regions and economic goods is a main research and application field of quantitative economics (Samuelson and Nordhaus, 2004). A widely used method for this purpose is the input-output analysis (IOA), developed by W. Leontief<sup>1</sup> (Leontief, 1986). This method uses input-output tables (IOTs) that express annual monetary input and output flows between economic industries within a region. An extension of these data sets are multi-regional input-output tables (MRIOTs), which include information about interregional trade flows. In particular, they describe annual monetary flows that leave an economic sector (e.g. Agriculture, Machinery, etc.) of a region (e.g. states, countries) and enter a target sector within a target region (Wiedmann et al., 2011). In the following, a sector  $i$  within a region  $r$  is called a regional sector  $ir$ . In addition to the intermediate flows between regional sectors, an MRIOT also contains information about the consumption of goods, expressed in monetary flows from regional sectors to the final demand of a region.

The column entries of an MRIOT represent inputs into the corresponding regional sector or region, while rows represent outputs of a regional sector. In general, MRIOTs may be supplemented with additional information, CO<sub>2</sub> emissions for instance. Thus, they often provide data for life cycle assessments (Algarin et al., 2015). The tables can be used for identifying supply chains, which can be defined as the procedure of processing supply material (inputs) into new products with a value added, which are purchased by and transported to customers, who in turn may further process them into more valuable goods. An exemplary supply chain could start with the mining sector in Australia, extracting ore that is purchased by the machinery sector of Japan, whose outputs are in turn consumed by German private

---

<sup>1</sup>Wassiliy Leontief (1906–1999), earned a Nobel Prize in Economics.

## 2 Data, methods and theory

region r	sector s	1				2				3				1			2			Total output
		1	2	3	4	1	2	3	4	1	2	3	4	r=1	r=2	r=3	r=1	r=2	r=3	
	index	ir=11	ir=21	ir=31	ir=41	ir=12	ir=22	ir=32	ir=42	ir=13	ir=23	ir=33	ir=43	r=1	r=2	r=3	r=1	r=2	r=3	Total output
1	1	346	156	95	594	819	154	832	397	409	562	241	554	394	902	446	6901			
1	2	354	443	7	908	42	92	561	839	470	770	83	368	514	694	512	6657			
1	3	291	795	243	825	753	2	340	232	251	605	526	610	384	753	909	7518			
1	4	637	259	289	813	500	716	947	645	856	221	898	41	91	653	301	7868			
2	1	547	466	910	276	518	149	779	553	197	285	305	828	630	565	857	7864			
2	2	752	936	822	638	611	496	98	924	608	689	872	972	847	209	37	9511			
2	3	295	444	7	828	929	535	367	257	890	429	641	26	165	419	886	7117			
2	4	113	518	791	459	79	748	254	218	586	673	424	157	800	355	501	6677			
3	1	46	457	552	572	632	680	730	607	796	186	15	958	338	320	194	7082			
3	2	962	96	544	96	675	113	711	337	787	571	241	211	479	14	608	6445			
3	3	531	190	686	191	374	615	788	738	351	32	565	622	269	814	559	7326			
3	4	857	776	897	18	915	482	308	458	253	145	982	270	700	822	729	8612			
Input s=1		939	1079	1557	1443	1969	983	2341	1557	1402	1032	561	2339	1361	1787	1497				
Input s=2		2068	1476	1373	1642	1327	701	1370	2100	1865	2030	1196	1551	1840	917	1157				
Input s=3		1116	1429	936	1845	2055	1152	1495	1227	1492	1065	1732	1258	818	1986	2354				
Input s=4		1607	1553	1976	1291	1493	1946	1509	1321	1695	1039	2305	468	1591	1830	1531				
Total input		5729	5537	5842	6220	6845	4782	6716	6206	6455	5167	5794	5617	5610	6520	6538				

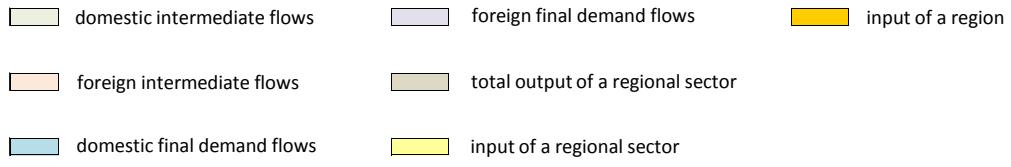


Figure 2.1: Exemplary MRIOT. Each row represents the output of a regional sector  $ir$  and its monetary flows to domestic regional sectors (green), foreign regional sectors (red), domestic final demand (blue) and foreign final demand (purple). The sum of these flows is the total output. The columns of the MRIOT represent the inputs of the regional sectors and the inputs of the final demand of each region. The total inputs of a regional sector or region are displayed at the input rows (yellow and orange).

households. Such a supply chain represents a directed monetary flow, as well as an exchange of information between supplier, processor and customer (Beamon, 1998).

An exemplary MRIOT is shown in Fig. 2.1. A range of selected MRIOT databases with relatively high regional and sectoral resolution are presented in Tab. 2.1.

### 2.1.2 The EORA world MRCIO database

An extensive MRCIO data set of high regional resolution is the EORA data set (Lenzen et al., 2012; Lenzen et al., 2013). The full data set comprises 15909 sectors and 186 regions, accompanied by various environmental and social satellite accounts. The developers also provide a sectorally aggregated EORA26<sup>2</sup> data set with 26 sectors<sup>3</sup> (see Fig. A.1) and 186 regions<sup>4</sup> (Fig. A.2). The monetary flows within the table are provided in basic or purchaser's prices respectively, annually and in units

<sup>2</sup>[www.worldmrio.com/simplified/](http://www.worldmrio.com/simplified/); accessed March 2016

<sup>3</sup>A more in-depth description of the sectors can be found at (United Nations, 2008)

<sup>4</sup>Since the EORA data is updated regularly, it is possible that new regions enter the data set. For that reason, the amount of regions used throughout this thesis (data download in 2013) may vary from that of the most recent version.

of 1000 US\$.

	Content	Regional resolution	Sectoral resolution	Years covered
EORA	Global MRIOT	186	15909	1990-2013
EORA26	Sectoral aggregated EORA	186	26	1990-2013
GTAP9	Global MRIOT	140	57	2004, 2007, 2011
WIOD	27 EU and 13 other countries	40	35	1995-2011
AIIOT	9 Asian countries plus USA	10	76	1985-2005
EXIOBASE	27 EU and 21 other countries	48	163	2000, 2007

Table 2.1: Overview of selected MRIOT databases (as of 2015): EORA (Lenzen et al., 2013), GTAP9 (Narayanan et al., 2015), WIOD (Timmer et al., 2015), AIIOT (IDE-JETRO, 2006), EXIOBASE (Tukker et al., 2009).

## 2.2 Methods: Linear Optimisation

### 2.2.1 Linear optimisation problems

A linear optimisation problem of standard form (Bertsimas and Tsitsiklis, 1997)

$$\begin{aligned} \min_{\mathbf{x}} \quad & F = \mathbf{c}^T \mathbf{x}, \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{2.1}$$

aims at minimising the target function  $F(\mathbf{x}) = \mathbf{c}^T \mathbf{x} : \mathbb{R}^n \rightarrow \mathbb{R}$ , with  $\mathbf{c} \in \mathbb{R}^n$  being the cost vector and  $\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \geq \mathbf{0}$  the nonnegative free variables whose values are constrained to the solution space  $X := \{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ , with  $\mathbf{A} \in \mathbb{R}^{m \times n}$  being the constraint matrix and  $\mathbf{b} \in \mathbb{R}^m$  the constraint vector. In particular,  $n$  free variables and  $n + m$  constraints exist. The above standard form is linear in the target function and in the constraints.

If  $\mathbf{A}$  is a square matrix  $m = n$  and a solution to  $\mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$  exists, it is unique. Otherwise, the constraint matrix  $\mathbf{A}$  is under-determined ( $m < n$ ), thus there may exist several feasible solution vectors  $\{\mathbf{x}\}$  satisfying the constraints. The solutions minimising the target function  $F(\mathbf{x})$  and satisfying the constraints form the set of optimal solutions  $\{\mathbf{x}^*\}$ .

## 2 Data, methods and theory

The set of feasible optimal solutions  $\{\mathbf{x}^*\}$  may consist of a number of elements specific to the optimisation problem. Possible results (see (Bertsimas and Tsitsiklis, 1997), p.24) are:

- A unique optimal solution  $\mathbf{x}^*$  exists
- Multiple optimal solutions  $\{\mathbf{x}^*\}$  exist
- The optimal target function is unbounded,  $F(\{\mathbf{x}^*\}) = -\infty$ , and therefore no optimal feasible solution exists.
- The set of feasible solutions is empty  $\{\mathbf{x}\} = \emptyset$ , thus no solution exists at all.

In linear programming, local optimality implies global optimality, because a convex function is minimized over a convex set (see (Bertsimas and Tsitsiklis, 1997), p.82).

To solve linear optimisation problems, appropriate algorithms are required. Two established methods are the simplex method and the interior point method, which are compared in Tab. 2.2. A key disadvantage of the interior point algorithm is the necessity of approximating gradients. In consideration of large problem dimensions, this may lead to computational performance losses due to the multivariable gradient calculations and may result in even less accurate approximated solutions. For these reasons, we assume the simplex method as the more suitable algorithm for this thesis. It is covered in detail in the next section.

	Simplex method	Interior point method
Method class	Basis exchange method	Gradient method
Geometrical interpretation	Moves along the edges of the polyhedron $X$	Moves through the interior of the polyhedron $X$
Solution accuracy	Exact solution	Approximate solution
Average case complexity	Linear in number of constraints: $\mathcal{O}(m)$	Logarithmic in number of variables: $\mathcal{O}(\log n)$
Worst case complexity	Exponential in number of variables: $\mathcal{O}(\beta^n)$	Polynomial in number of variables: $\mathcal{O}(n^\beta)$

Table 2.2: Comparison of properties of the simplex and interior point method (arbitrary  $\beta \in \mathbb{R}^+$ ) (Bertsimas and Tsitsiklis, 1997).

## 2.2.2 Simplex algorithm

The simplex method - developed by George Dantzig in 1947 (Dantzig, 1987) - is applied to solve linear programs of the standard form (Eq. (2.1)). To understand the functionality of the algorithm, a geometric perspective is helpful. In the next paragraphs, we develop geometric properties of the optimisation problem and its solution space  $X$  and identify them with algebraic properties that are hereafter used in the simplex application scheme.

### Geometric concept

The solution space  $X$  is restricted by hyperplanes  $\mathbf{A}\mathbf{x} = \mathbf{b}$  and halfspaces  $\mathbf{x} \geq \mathbf{0}$ , forming the shape of a convex polyhedron (Press et al., 2002). Here we assume that the rows of matrix  $\mathbf{A}$  are linearly independent (full rank).<sup>5</sup> For better understanding, an exemplary linear problem and the corresponding geometrical view is shown in Fig.2.2. The target function is linear, meaning if we only change one component  $x_i$  and leave the others constant,  $F$  will change linearly with  $x_i$ . Analogously, if the value of  $F$  is assumed to be constant  $F = k$ , all possible variable combinations fulfilling this condition are linearly dependent from each other and therefore create hyperplanes inside the solution space (see Fig. 2.3). Hyperplanes with  $F = q \neq k$  are parallel to  $F = k$ . To find the optimal value of  $F$ , the target hyperplane has

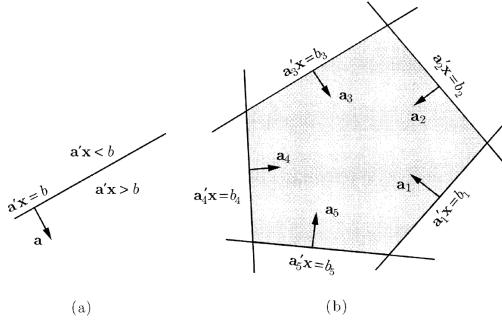


Figure 2.2: Geometrical properties of a linear optimisation problem. (a) shows a hyperplane  $\{\mathbf{x} \mid \mathbf{a}\mathbf{x} = \mathbf{b}\}$  and two halfspaces  $\{\mathbf{x} \mid \mathbf{a}\mathbf{x} \geq \mathbf{b}\}$ . The vector  $\mathbf{a}$  is a row vector of the matrix  $\mathbf{A}$ , representing one constraint. (b) View of the polyhedron solution space  $X = \{\mathbf{x} \mid \mathbf{a}_i\mathbf{x} = b_i, i = 1, \dots, 5\}$ . The bold lines represent the hyperplanes  $\{\mathbf{x} \mid \mathbf{a}_i\mathbf{x} = b_i, i = 1, \dots, 5\}$ . Since the solution space contains five halfspaces, the whole shaded part inside the convex polyhedron is the feasible solution space  $X$  (Image taken from (Bertsimas and Tsitsiklis, 1997), p.44).

<sup>5</sup>If not so, linearly dependent rows or columns can be removed from the matrix to make it full rank.

## 2 Data, methods and theory

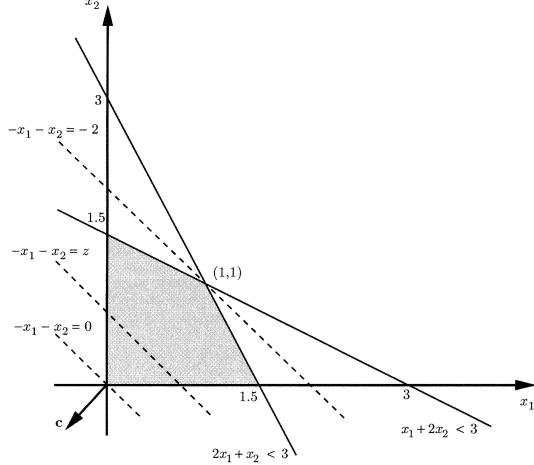


Figure 2.3: Geometrical view of the solution space  $X$  of an exemplary linear program with  $F = -x_1 - x_2$  and  $X = \{\mathbf{x} \mid x_1 + 2x_2 \leq 3; x_1 + 2x_2 \leq 3; x_1, x_2 \geq 0\}$ . The optimum of the target function has to lie on a vertex (corner) of the convex solution space (shaded area), in this case at  $(x_1, x_2) = (0, 0)$  (Image taken from (Bertsimas and Tsitsiklis, 1997), p.22).

to touch one of the vertices of  $X$ , being parallel to all other hyperplanes. If several vertices represent the same minimal value of  $F$ , the optimal solution lies on the edge of the polyhedron and therefore is not unique. Altogether, optimal solutions lie on the vertices of the polyhedron. With a given vector  $\mathbf{x}^1 \in \mathbb{R}^n$  (not necessarily a solution) that satisfies the  $i$ -th constraint  $\sum_j A_{ij}x_j^1 = b_i$ , the constraint  $i$  is said to be active at  $\mathbf{x}^1$ . Assuming  $\mathbf{x}^1$  satisfies  $n$  constraints of all  $m + n$  constraints, including all equality constraints, and all of them are linearly independent, then  $\mathbf{x}^1$  is called a basic solution of the full system of equations. Still, there may be inactive constraints that  $\mathbf{x}^1$  does not satisfy. In this case  $\mathbf{x}^1$  is a basic but not necessarily a basic feasible solution.

If  $\mathbf{x}^1$  is a basic solution and satisfying all  $m + n$  constraints, then it is a basic and feasible solution  $\bar{\mathbf{x}}^1$ . A basic feasible solution can be identified as a corner of the polyhedron and according to the *fundamental theorem of linear programming* (Bertsimas and Tsitsiklis, 1997), the optimal solution  $\bar{\mathbf{x}}^*$  can be found at a corner of the polyhedron and is therefore also a basic feasible solution.

Two basic feasible solutions  $\bar{\mathbf{x}}^1, \bar{\mathbf{x}}^2$  are adjacent to one another, if there are  $n - 1$  linearly independent constraints that are active at both points. In the geometrical representation, two corner solutions are adjacent if the line that connects both of them can be identified as an edge of the polyhedron.

To find a basic feasible solution in the case of the standard form optimisation problem, at first all equality constraints of  $\mathbf{A}\mathbf{x} = \mathbf{b}$  are chosen as active constraints and  $m$  components of  $\mathbf{x}$  are chosen to be basic components. Furthermore one chooses  $n - m$  of the remaining nonnegativity constraints as active by assigning zero to the corresponding components of  $\mathbf{x}$ , which are called non-basic components. This gives a full set of  $n$  active constraints and thus a basic solution. Ensuring that the choice of  $\mathbf{x}$  does not violate the inactive constraints, the basic solution is a basic feasible solution  $\bar{\mathbf{x}}$ .

### Algebraic concept

The fundamental concept of the simplex method is to start from one basic feasible solution  $\bar{\mathbf{x}}^1$  in the corner of the polyhedron and move along one selected edge towards another corner  $\bar{\mathbf{x}}^2$ , meanwhile reducing the target function. Repetition of this process leads to the optimal solution  $\bar{\mathbf{x}}^*$ . This procedure of changing the basic solutions iteratively is called a basis exchange method.

As mentioned above, a basic solution  $\bar{\mathbf{x}} \in X$  can be expressed through  $m$  basic components  $\bar{\mathbf{x}}_J \geq \mathbf{0}$ , assigned by indices  $\{j\} = J \subseteq N := \{1, \dots, n\}$  and  $n - m$  non-basic components  $\bar{\mathbf{x}}_K = \mathbf{0}$ , identified by indices  $\{k\} = K = N \setminus J$ , so that  $\bar{\mathbf{x}} = (\bar{\mathbf{x}}_J, \bar{\mathbf{x}}_K) = (\bar{\mathbf{x}}_J, \mathbf{0})$ . Analogously, the constraint matrix  $\mathbf{A}$  can be separated into a submatrix  $\mathbf{A}_J = (\mathbf{a}_{1j}, \dots, \mathbf{a}_{mj})^T \in \mathbb{R}^{m \times m}$ ,  $\forall j \in J$ , containing the columns corresponding to the active components, and a matrix  $\mathbf{A}_K = (\mathbf{a}_{1k}, \dots, \mathbf{a}_{mk})^T \in \mathbb{R}^{m \times n-m}$ ,  $\forall k \in K$ , containing the inactive columns of  $\mathbf{A}$ , with  $\mathbf{A} = (\mathbf{A}_J, \mathbf{A}_K)$ . A basic solution  $\bar{\mathbf{x}}$  always comes along with a corresponding basic matrix  $\mathbf{A}_J$ . Using the above partitions yields

$$\mathbf{b} = \mathbf{A}\bar{\mathbf{x}} = \mathbf{A}_J\bar{\mathbf{x}}_J + \mathbf{A}_K\bar{\mathbf{x}}_K , \quad (2.2)$$

$$\xrightarrow{\bar{\mathbf{x}}_K=0} \mathbf{b} = \mathbf{A}_J\bar{\mathbf{x}}_J . \quad (2.3)$$

With knowledge of a basic solution  $\bar{\mathbf{x}} = (\bar{\mathbf{x}}_J, \bar{\mathbf{x}}_K)$  other arbitrary solutions  $\mathbf{x} = (\mathbf{x}_J, \mathbf{x}_K) \in X$  ( $\mathbf{x}_K$  not necessarily zero) on the edge of the polyhedron can be created, using Eq. (2.3)

$$\mathbf{b} = \mathbf{A}\mathbf{x} = \mathbf{A}_J\mathbf{x}_J + \mathbf{A}_K\mathbf{x}_K \quad (2.4)$$

$$\leftrightarrow \mathbf{x}_J = \mathbf{A}_J^{-1}\mathbf{b} - \mathbf{A}_J^{-1}\mathbf{A}_K\mathbf{x}_K \quad (2.5)$$

$$= \bar{\mathbf{x}}_J - \mathbf{A}_J^{-1}\mathbf{A}_K\mathbf{x}_K . \quad (2.6)$$

## 2 Data, methods and theory

Substituting  $\mathbf{x}_K$  with the parameter  $\boldsymbol{\lambda}_K$  leads to

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_J \\ \mathbf{x}_K \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{x}}_J \\ \mathbf{0}_K \end{pmatrix} + \begin{pmatrix} -\mathbf{A}_J^{-1}\mathbf{A}_K \\ \mathbf{I}_{n-m} \end{pmatrix} \boldsymbol{\lambda}_K = \bar{\mathbf{x}} - \mathbf{W}_K \boldsymbol{\lambda}_K \geq \mathbf{0}, \quad (2.7)$$

with

$$\begin{pmatrix} \mathbf{A}_J^{-1}\mathbf{A}_K \\ -\mathbf{I}_{n-m} \end{pmatrix} = \begin{pmatrix} \mathbf{W}_K^{(J)} \\ \mathbf{W}_K^{(K)} \end{pmatrix} = \mathbf{W}_K \in \mathbb{R}^{n \times (n-m)}. \quad (2.8)$$

By increasing the value of  $\lambda_l, l \in K$ , one moves along a ray  $\mathbf{x}(\lambda_l)$  on the surface of the solution space  $X$ , originating in  $\bar{\mathbf{x}}$ . The whole set of rays  $\mathbf{x}(\{\lambda_l\})$  starting at  $\bar{\mathbf{x}}$  creates a cone (a simplex), which gives the algorithm its name. One single elementary ray is

$$\mathbf{x}(\lambda_l) := \bar{\mathbf{x}} - \lambda_l \mathbf{w}_l \leftrightarrow \begin{cases} \mathbf{x}_J(\lambda_l) = \bar{\mathbf{x}}_J - \lambda_l \mathbf{A}_J^{-1} \mathbf{a}_l \\ x_k(\lambda_l) = \lambda_l \delta_{kl}, k \in K \end{cases}, \quad (2.9)$$

with  $\mathbf{w}_l$  being the l-th column of  $\mathbf{W}_K$ ,  $\mathbf{a}_l$  the l-th column of  $\mathbf{A}_K$  and  $\delta_{kl}$  being the Kronecker Delta.

To determine which ray originating in  $\bar{\mathbf{x}}$  reduces the target function, we consider the cost function

$$\mathbf{c}^T \mathbf{x} = \mathbf{c}_J^T \mathbf{x}_J + \mathbf{c}_K^T \mathbf{x}_K = \mathbf{c}_J^T (\bar{\mathbf{x}}_J - \mathbf{A}_J^{-1} \mathbf{A}_K \mathbf{x}_K) + \mathbf{c}_K^T \mathbf{x}_K \quad (2.10)$$

$$= \underbrace{\mathbf{c}_J^T \bar{\mathbf{x}}_J}_{\mathbf{c}^T \bar{\mathbf{x}}} + \underbrace{(\mathbf{c}_K^T - \mathbf{c}_J^T \mathbf{A}_J^{-1} \mathbf{A}_K)}_{\gamma_K^T} \mathbf{x}_K = \mathbf{c}^T \bar{\mathbf{x}} + \gamma_K^T \mathbf{x}_K. \quad (2.11)$$

The cost function in the area of  $\bar{\mathbf{x}}$  changes by increasing the non-basis variables  $\mathbf{x}_K = \boldsymbol{\lambda}_K$ , weighted by the vector of reduced costs  $\gamma_K$ . In each iteration, moving along a ray  $\mathbf{x}(\lambda_l)$  with negative reduced costs  $\gamma_l$  leads towards the minimal solution of the linear problem. The minimum is reached when each element of  $\gamma_K$  is non-negative  $\gamma_l \geq 0, \forall l$ . In this case, the basic corner solution of the current iteration  $\bar{\mathbf{x}}$  is the optimal solution  $\bar{\mathbf{x}}^*$ .

**Theorem 1** Let  $\bar{\mathbf{x}}$  be a basic feasible solution and  $\gamma_K$  the corresponding vector of reduced costs.

If  $\gamma_l \geq 0, \forall l \in K$ , then  $\bar{\mathbf{x}} = \bar{\mathbf{x}}^*$  is the optimal solution.

## 2.2 Methods: Linear Optimisation

While moving along the ray it is important to ensure that every component of  $\mathbf{x}$  remains inside the solution area  $X$ . Consider ray  $\mathbf{x}(\lambda_l)$  in component form

$$x_j(\lambda_l) = \bar{x}_j - \lambda_l w_{jl} . \quad (2.12)$$

While increasing  $\lambda_l \geq 0$ , the components  $x_j(\lambda_l)$  increase if  $w_{jl} \leq 0$ . For executing a basis exchange, a parameter  $\lambda_l$  with  $\gamma_l < 0$  is increased until one component  $x_p(\lambda_l)$  gets equal to zero

$$\lambda_l := \min \left\{ \frac{\bar{x}_j}{w_{jl}} : j \in J, w_{jl} > 0 \right\} = \frac{\bar{x}_p}{w_{pl}} \geq 0 , \quad (2.13)$$

which is afterwards moved from the vector of basic components  $\bar{\mathbf{x}}_J$  into the new vector of non-basic components  $\bar{\mathbf{x}}_{K'}$ . Only those components with  $w_{jl} > 0$  are taken into account. If  $w_{jl} \leq 0, \forall j \in J$ , the solution is unbounded. The index  $p$  is added to the new inactive set  $K'$ , while variable  $l$  is included into the new basis set  $J'$

$$J' = J \setminus \{p\} \cup \{l\} , \quad (2.14)$$

$$K' = K \setminus \{l\} \cup \{p\} , \quad (2.15)$$

resulting in a new corner solution  $\bar{\mathbf{x}}' = (\bar{\mathbf{x}}_{J'}, \bar{\mathbf{x}}_{K'})$  that brings the target function  $F$  closer to the minimum. If the new target function value happens to be  $F = -\infty$  though, no optimal feasible solution exists.

An overview about the simplex algorithm application scheme is given in Tab. 2.3.

Input	Initial feasible basis $\mathbf{A}_J$ , $J \subseteq N$ , $N := \{1, \dots, n\}$
Step 1	$\bar{\mathbf{x}}_J := \mathbf{A}_J^{-1} \mathbf{b}$ , $K := N \setminus J$
Step 2	if $F(\bar{\mathbf{x}}) = -\infty$ : <b>STOP, no optimal solution</b>
Step 3	search for $\gamma_l < 0, l \in K$ out of $\gamma_K := \mathbf{c}_K^T - \mathbf{c}_J^T \mathbf{A}_J^{-1} \mathbf{A}_K$
Step 4	if $\gamma_l \geq 0 \ \forall l \in K$ : <b>STOP, optimum found</b>
Step 5	$\mathbf{w}_l^{(J)} := \mathbf{A}_J^{-1} \mathbf{A}_l$
Step 6	if $w_{il} \leq 0 \ \forall i \in J$ : <b>STOP, unbounded solution</b>
Step 7	Identify $p \in J$ with: $\frac{\bar{x}_p}{w_{pl}} := \min \left\{ \frac{\bar{x}_i}{w_{il}} : w_{il} > 0, i \in J \right\} = \lambda_l$
Step 8	$J := J \setminus \{p\} \cup \{l\}$ ; <b>jump to step 1</b>

Table 2.3: Simplex application scheme (phase II).

### Simplex phase I

The simplex application scheme starts with an initial feasible basis  $\mathbf{A}_J$ , which is not always trivial to obtain without applying a so called “phase I” procedure. Again, we start with the standard form of optimisation problems

$$\begin{aligned} \min_x \quad & F = \mathbf{c}^T \mathbf{x}, \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0}. \end{aligned} \tag{2.16}$$

In addition to the free variables  $\mathbf{x} \in \mathbb{R}^n$ , new artificial slack variables  $\mathbf{y} \in \mathbb{R}^m$  are introduced that transform the optimisation problem into a new auxiliary problem (see (Bertsimas and Tsitsiklis, 1997), p.111)

$$\begin{aligned} \min_x \quad & \tilde{F} = \tilde{\mathbf{c}}^T \mathbf{y}, \\ \text{s.t.} \quad & \mathbf{Ax} + \mathbf{y} = \mathbf{b}, \\ & \mathbf{x}, \mathbf{y} \geq \mathbf{0}. \end{aligned} \tag{2.17}$$

Those constraints with  $b_i < 0$  of the original problem are multiplied by  $-1$ , ensuring that  $b_i \geq 0, \forall i$ . It is easy to find a basic feasible solution for the auxiliary problem by setting  $\mathbf{x} = \mathbf{0}$  and  $\mathbf{y} = \mathbf{b}$  and  $\mathbf{A}_J = \mathbf{1}^{m \times m}$ .

Input	Start with auxiliary problem. Multiply constraints with $b_i < 0$ with $-1$ . Starting basis: $\mathbf{A}_J = \mathbf{1}^{m \times m}$ , $\bar{\mathbf{x}}_J = \mathbf{y} = \mathbf{b}$
Step 1	Apply the simplex method to the auxiliary problem with cost $\tilde{F}(\mathbf{z}) = \sum_{i=1}^m y_i$ . The free variables are $\mathbf{z} = (\mathbf{x}, \mathbf{y})$ and the optimal solution $\bar{\mathbf{z}}^*$ .
Step 2	If $\tilde{F}(\bar{\mathbf{z}}^*) > 0$ : <b>STOP, the original problem is infeasible</b>
Step 3	If $\tilde{F}(\bar{\mathbf{z}}^*) = 0$ : Remove the artificial variables and return to the original problem.
Step 4	Let the final basis be the initial basis $\mathbf{A}_J$ for simplex Phase II.

Table 2.4: Simplex application scheme (phase I).

By applying the simplex scheme to the auxiliary problem, the artificial slack variables  $\mathbf{y}$  are driven out of the basis, since they appear in the target function, while the original free variables  $\mathbf{x}$  do not. The herewith created basis can now be

used as an initial basis of the original problem and the optimum can be obtained by applying the standard simplex algorithm (simplex phase II). If not all auxiliary variables are driven out of the basis while applying the phase I algorithm, the original problem is not feasible. The phase I algorithm thus also provides a feasibility test.

An overview about the simplex algorithm phase I application scheme is given in Tab. 2.4

### 2.2.3 Absolute value target function

Consider the following optimisation problem with nonnegative cost coefficients  $c_i$  minimising the summed absolute values of  $x_i$

$$\begin{aligned} & \min \sum_i c_i |x_i|, \\ & \text{s.t. } \sum_i A_{ji} x_i = b_j, \forall j, \\ & \quad x_i \geq 0, \forall i. \end{aligned} \tag{2.18}$$

This is a piecewise linear convex problem since the absolute value target function has summands that are piecewise linear convex functions (due to the nonnegative  $c_i$ ). The target function can be transformed into a full linear one by substituting  $x_i$  with two new non-negative variables (see (Bertsimas and Tsitsiklis, 1997), p.17):

$$x_i = x_i^+ - x_i^-, \tag{2.19}$$

$$x_i^+ \geq 0, x_i^- \geq 0. \tag{2.20}$$

If  $x_i^+ > x_i^-$ , then  $x_i$  takes on a positive value. Conversely, it gets negative with  $x_i^- > x_i^+$ . The reformulated optimisation problem is

$$\begin{aligned} & \min \sum_i c_i |x_i^+ - x_i^-|, \\ & \text{s.t. } \sum_i A_{ji} (x_i^+ - x_i^-) = b_j, \forall j, \\ & \quad x_i^+ \geq 0, x_i^- \geq 0, \forall i. \end{aligned} \tag{2.21}$$

Still, the problem is not fully linear, but if either  $x_i^+$  or  $x_i^-$  equals zero (hence the product of both equals zero), the absolute value function can be rewritten as two

## 2 Data, methods and theory

new absolute value terms

$$x_i^+ \cdot x_i^- = 0, \quad (2.22)$$

$$\rightarrow |x_i^+ - x_i^-| = |x_i^+| + |x_i^-| = x_i^+ + x_i^-, \quad (2.23)$$

leading to the linear problem

$$\begin{aligned} & \min \sum_i c_i(x_i^+ + x_i^-), \\ & \text{s.t. } \sum_i A_{ji} (x_i^+ - x_i^-) = b_j, \forall j, \\ & x_i^+ \cdot x_i^- = 0, \forall i, \\ & x_i^+ \geq 0, x_i^- \geq 0, \forall i. \end{aligned} \quad (2.24)$$

The constraint on the product of  $x_i^+$  and  $x_i^-$  can be dropped, because minimisation of the target function will automatically lead to one of the two variables being zero. If it would not, one could reduce both by the same value, what keeps the solution feasible and reduces the target function's value. Thus, the nonlinear constraint  $x_i^+ \cdot x_i^- = 0$  is always fulfilled when using the above target function and the final problem formulation becomes linear:

$$\begin{aligned} & \min \sum_i c_i(x_i^+ + x_i^-), \\ & \text{s.t. } \sum_i A_{ji} (x_i^+ - x_i^-) = b_j, \forall j, \\ & x_i^+ \geq 0, x_i^- \geq 0, \forall i. \end{aligned} \quad (2.25)$$

Altogether, this technique adds one more variable per initial free variable to the original linear problem.

### 2.3 Theory: Network properties

In the following, network properties of relevance for this work are presented.

#### Network definition

A network  $G$  is defined as  $G = (V, \mathbf{E})$ , where the set  $V \in \mathbb{R}^v$  contains  $v$  elements that represent nodes. The matrix  $\mathbf{E} \in \{0, 1\}^{v \times v}$  is called adjacency matrix

(Newman, 2010), with

$$E_{ij} = \begin{cases} 1 & \text{if a connection between node } i \text{ and } j \text{ exists,} \\ 0 & \text{otherwise} \end{cases}, \forall i, j \in V. \quad (2.26)$$

If the adjacency matrix  $\mathbf{E}$  is symmetric, the corresponding network is undirected, whereas an unsymmetric matrix indicates a directed network, where edges have a distinct source and target node.

### Weight of an edge

Complementary to the adjacency matrix, a weight matrix  $\mathbf{W} \in \mathbb{R}^{v \times v}$  can be defined, assigning a weight  $w_{ij}$  to each existing edge  $E_{ij} = 1$  between two nodes  $i$  and  $j$ .

### Flow networks

A flow network or transportation network is a directed graph with weighted edges. The weights represent units of a property being transmitted through the network. Each node  $i$  receives input flows  $w_{ji}$  and distributes them as output flows  $w_{ij}$ . A node whose inputs do not equal its outputs is either a source node with a higher output  $X_i$  than input  $D_i$ , or a sink node, with higher inputs than outputs.

$$X_i = \sum_j^v W_{ij}, \quad (2.27)$$

$$D_i = \sum_j^v W_{ji}. \quad (2.28)$$

### Degree of a node

The degree is the amount of a node's connections (Newman, 2010). In the case of a directed network, one can distinguish between the incoming degree  $\deg_i^{in}$  and outgoing degree  $\deg_i^{out}$

$$\deg_i^{in} = \sum_j^v E_{ji}, \quad (2.29)$$

$$\deg_i^{out} = \sum_j^v E_{ij}, \quad (2.30)$$

$$\deg_i = \deg_i^{out} + \deg_i^{in} \quad (2.31)$$

$$(2.32)$$

## 2 Data, methods and theory

### Shared neighbours

The set of neighbours of a node  $i$  contains the nodes directly connected to it through an edge. In the case of a directed graph, one distinguishes between the set of neighbours  $N_i^{out}$  connected by an outgoing edge from  $i$ , and the set  $N_i^{in}$  connected by incoming edges into  $i$

$$N_i^{in} = \{j \mid E_{ji} = 1; j \neq i \in V\}, \quad (2.33)$$

$$N_i^{out} = \{j \mid E_{ij} = 1; j \neq i \in V\}, \quad (2.34)$$

$$N_i = N_i^{out} + N_i^{in} \quad (2.35)$$

$$= \{j \mid E_{ji} = 1 \vee E_{ij} = 1; j \neq i \in V\}. \quad (2.36)$$

The set of shared neighbours  $N_{ij}$  of two nodes  $i$  and  $j$  is the intersection of the two individual neighbour sets

$$N_{ij}^{in} = N_i^{in} \cap N_j^{in}, \quad (2.37)$$

$$N_{ij}^{out} = N_i^{out} \cap N_j^{out}, \quad (2.38)$$

$$N_{ij} = N_i \cap N_j. \quad (2.39)$$

# 3 Problem description

## 3.1 Optimal trade network response to production failure

### Network representation of the EORA data

A representation of the global trade network can be built on a multi-regional input-output table  $\mathbf{M} = (\mathbf{Z}, \mathbf{Y})$ , consisting of the intermediate flow matrix  $\mathbf{Z} \in \mathbb{R}_+^{n_{rs} \times n_{rs}}$  and the final demand flow matrix  $\mathbf{Y} \in \mathbb{R}_+^{n_{rs} \times n_r}$  (Bierkandt et al., 2014; Maluck and Donner, 2015). The data can be used to form a network with  $n_r$  regions, each containing  $n_s$  economic sectors, so that  $n_{rs} = n_r \cdot n_s$  yields the total amount of regional sectors. A sector  $i$  in region  $r$  is called regional sector  $ir$ . All regions and regional sectors can be interpreted as nodes of the network, which are connected by the intermediate flows  $Z_{ir \rightarrow js}$  between two regional sectors  $ir$  and  $js$  and the final demand flows  $Y_{ir \rightarrow s}$  between a regional sector  $ir$  and a region  $s$ . In the whole network  $n_Z$  intermediate flows and  $n_Y$  final demand flows exist. Both flows are interpreted as weighted edges between the nodes and are directly taken from the corresponding MRIOT cell entries. The total production of a regional sector is given by  $X_{ir} = \sum_{js} Z_{ir \rightarrow js} + \sum_s Y_{ir \rightarrow s}$ . An exemplary translation of an MRIOT to a network view is given in Fig. 3.1. The network formed of the EORA data is visualised in Fig. 3.2. It is a directed flow network, where the nodes act, dependent on their inputs, as sources and the final demand acts as a sink.

Further theoretical analysis of the EORA data was conducted from a network of networks perspective (Maluck and Donner, 2015).

### Response to exogenous forcing

Extreme weather events can be interpreted as external forcings  $\epsilon$  reducing the production  $X_{i^*r^*}$  of a specific regional sector  $i^*r^*$ , that would lead to a lack of inputs for nodes further down the supply chain and an overproduction for nodes up the chain. We seek an optimal response of all regional sectors' productions and trade flows under which the final demand of the countries is still satisfied, such that the

### 3 Problem description

sum of absolute deviations  $\sum_{ir} |p_{ir} - 1|$  of the relative production changes

$$p_{ir} = \frac{X_{ir}^{\text{new}}}{X_{ir}^{\text{old}}} \quad (3.1)$$

is minimal. This gives a linear optimisation problem which can be solved e.g. by the simplex algorithm.

While the weights of network links can change, a restructuring is not possible, so the topology is not changed. Furthermore, we do not assume storage facilities, so the whole production of one node has to be transported to other nodes. Moreover, we take the *Leontief production function* into account, which indicates that the amount of a regional sector's supplies is proportional to the amount of produced goods and the proportions between different input goods are fixed (input goods cannot be substituted by input goods of another sector) (Allen, 1968).

region	sector		r1	r1	r2	r2	r3	r3	r1	r2	r3	Total output
			s1	s2	s1	s2	s1	s2	FD1	FD2	FD3	
	index		ir=11	ir=21	ir=12	ir=22	ir=13	ir=23	r1	r2	r3	
r1	s1	ir=11	10					10	10			30
	s2	ir=21	10	10					10			30
r2	s1	ir=12			10				10	10		30
	s2	ir=22			10				10	10	10	40
r3	s1	ir=13					10			10		20
	s2	ir=23					10			20		30
Input s1		10			10		10		20	10	10	
Input s2		10	10		10		10		20	10	30	
Total input		20	10	10	10	10	20	10	40	20	40	

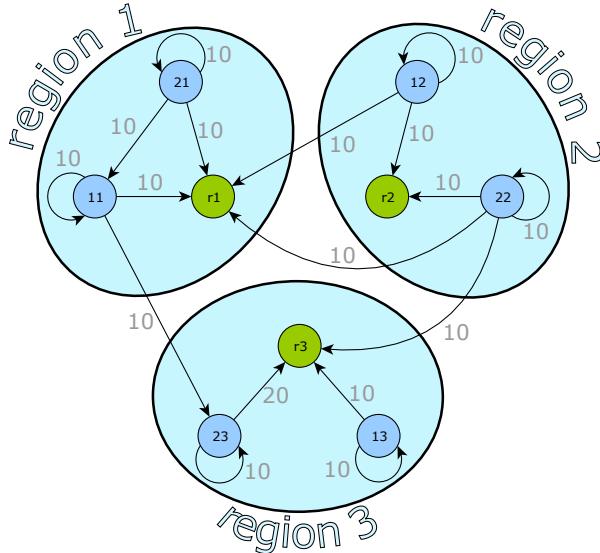


Figure 3.1: Network representation of MRIO data for a sample data set with  $n_r = 3$  regions,  $n_s = 2$  sectors and  $n_{rs} = 6$  regional sectors. Each region consists of two sectors (blue nodes) and a consumption site (green nodes).

### 3.1 Optimal trade network response to production failure

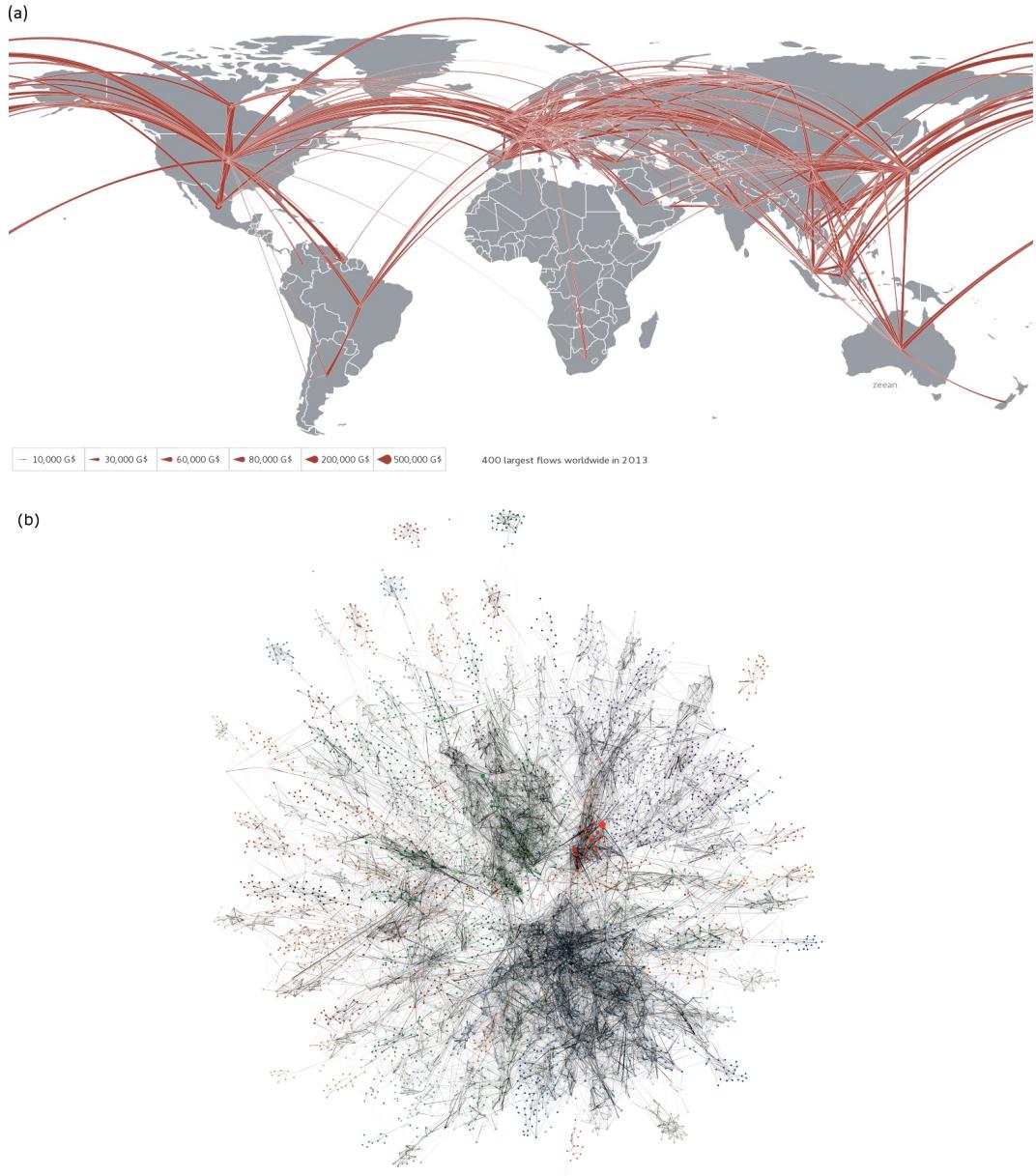


Figure 3.2: Visualisations of the EORA26 network. (a) Demonstration of the 400 largest monetary flows between countries for the year 2013 (taken from zeean.net). (b) A force-directed graph view of the network. Three big clusters are noticeable: The upper left one contains Asian regional sectors (green), the upper right one North American nodes (red) and the lower one is mostly made of European regional sectors (blue). The size of the nodes represents their amount of production and consumption. Many small clusters of regional sectors are found on the edge of the network, representing domestic economies without strong interrelations to other countries, which are often African nodes (orange). (Picture by S. Willner, 2015)



## 4 Model setup

### 4.1 Formulation of the linear optimisation problem

The model used to simulate a possible response of the world trade network to exogenous forcings  $\epsilon$  is set up as a linear optimisation problem, with the EORA26 MRIOT as input data. One selected regional sector  $i^*r^*$  is assumed to be affected by an exogenous forcing  $\epsilon$  and reduces its production according to

$$X_{i^*r^*}^1 = \hat{p} X_{i^*r^*}^0 \quad (4.1)$$

$$= (1 - \epsilon) X_{i^*r^*}^0, \quad (4.2)$$

with  $\hat{p} = (1 - \epsilon)$  being the final production ratio of the affected regional sector  $i^*r^*$ , fixed by the exogenous forcing and  $X_{i^*r^*}^0$ ,  $X_{i^*r^*}^1$  being the original production and production in the new state respectively. The response of the network to the disruption is governed by the target function and the constraints of the linear problem.

#### 4.1.1 Free variables and target function

The free variables of the optimisation problem  $\mathbf{x} = (\mathbf{p}, \mathbf{q}^Z, \mathbf{q}^Y)$  are the variables of relative production change  $\mathbf{p} = (p_{ir}) \in \mathbb{R}^{n_{rs}}$  for each regional sector and the variables of relative intermediate flow change  $\mathbf{q}^Z = (q_{ir \rightarrow js}) \in \mathbb{R}^{n_Z}$  and relative final demand flow change  $\mathbf{q}^Y = (q_{ir \rightarrow s}) \in \mathbb{R}^{n_Y}$  between the initial and the adapted network:

$$p_{ir} = \frac{X_{ir}^1}{X_{ir}^0}, \quad (4.3)$$

$$q_{ir \rightarrow js} = \frac{Z_{ir \rightarrow js}^1}{Z_{ir \rightarrow js}^0}, \quad (4.4)$$

$$q_{ir \rightarrow s} = \frac{Y_{ir \rightarrow s}^1}{Y_{ir \rightarrow s}^0}, \quad (4.5)$$

where  $p_{ir} = 1$  holds for the original state of the network as provided by the EORA data. Relative changes (positive or negative) of free variables are penalised by the

## 4 Model setup

target function

$$F(\mathbf{x}) = \sum_k c_k |x_k - 1|. \quad (\text{T1})$$

In the following, the target function is expressed as 'network adaptation effort', indicating the effort the network has to expend in order to adapt to the failure. The cost vector is set to

$$\mathbf{c} = (\underbrace{1, \dots, 1}_{n_{rs}}, \underbrace{0, \dots, 0}_{n_Y + n_Z}), \quad (4.6)$$

so deviations of the production ratio from  $p_{ir} = 1$  are penalised with costs of 1, while the cost for flow adaptation is set to 0, assuming that changes of production cause significantly higher effort than changes of transportation, which can therefore be neglected and approximated with zero cost. The cost vector reduces the network adaptation effort  $F$  to

$$F(\mathbf{p}) = \sum_{ir} |p_{ir} - 1|. \quad (\text{T1})$$

In addition to the free variables and target function, the constraints  $\mathbf{Ax} = \mathbf{b}$  are defined in the next two chapters. Two alternative constraint formulations are given: The general linear problem (LPG) guarantees the highest degree of adaptability, since all variables  $\mathbf{x}$  are considered as free variables. Contrarily, the special linear problem (LPS) puts a focus on the disrupted sector  $i^*$  and reduces the set of free variables to flows of these goods.

### 4.1.2 General linear problem (LPG)

$$\min_{p_{ir}} F = \sum_{ir} |p_{ir} - 1|, \quad (\text{T1})$$

$$s.t. \quad \sum_r q_{ir \rightarrow s} Y_{ir \rightarrow s} = \sum_r Y_{ir \rightarrow s}, \forall s, \forall i, \quad (\text{C1})$$

$$\sum_r q_{ir \rightarrow js} Z_{ir \rightarrow js} = p_{js} \sum_r Z_{ir \rightarrow js}, \forall js, \forall i, \quad (\text{C2})$$

$$p_{ir} X_{ir} = \sum_{js} q_{ir \rightarrow js} Z_{ir \rightarrow js} + \sum_s q_{ir \rightarrow s} Y_{ir \rightarrow s}, \forall ir, \quad (\text{C3})$$

$$p_{i^*r^*} = \hat{p}. \quad (\text{C4})$$

## 4.1 Formulation of the linear optimisation problem

This linear problem formulation is composed of the target function (T1) and four constraints (C1), (C2) (C3) and (C4). All production ratio variables  $p_{ir}$  and flow ratio variables  $q_{ir \rightarrow js}$ ,  $q_{ir \rightarrow s}$  are treated as free variables, so the solution vector  $\mathbf{x}$  yields

$$\mathbf{x} = (\mathbf{p}, \mathbf{q}^Z, \mathbf{q}^Y). \quad (4.7)$$

### Final demand constraint

The constant final demand of each region continues to be satisfied in the disrupted network state:

$$\sum_r q_{ir \rightarrow s} Y_{ir \rightarrow s} = \sum_r Y_{ir \rightarrow s}, \forall s, \forall i. \quad (C1)$$

The equation's left-hand side sums up all adaptable final demand flows  $q_{ir \rightarrow s} Y_{ir \rightarrow s}$  of a sector  $i$  into a region  $s$ . The result has to equal the sum of all final demand flows  $Y_{ir \rightarrow s}$  of the initial unperturbed network, so the final demand can still be fulfilled. Each region receives at most  $n_s$  different final demand input goods, therefore the maximum number of final demand constraints of the linear problem amounts to  $n_{C1} = n_r \cdot n_s$  constraints.

### Supply scaling constraint

Each regional sector receives sufficient inputs to produce the output of the new network state:

$$\sum_r q_{ir \rightarrow js} Z_{ir \rightarrow js} = p_{js} \sum_r Z_{ir \rightarrow js}, \forall js, \forall i. \quad (C2)$$

The sum of all adaptable intermediate input flows  $q_{ir \rightarrow js} Z_{ir \rightarrow js}$  of a sector  $i$  into a node  $js$  has to match the sum of its original inputs  $Z_{ir \rightarrow js}$ , scaled by its adaptable production ratio  $p_{js}$ .<sup>1</sup>

In a fully connected network, each node  $js$  requires an input of each type of good, hence the maximum number of supply scaling constraints of the linear problem amounts to  $n_{C2} = n_{rs} \cdot n_s = n_s^2 \cdot n_r$ .

### Production output balance constraint

The total output of a regional sector is completely distributed to other network

---

<sup>1</sup>and therefore satisfying the Leontief production function.

## 4 Model setup

nodes:

$$p_{ir}X_{ir} = \sum_{js} q_{ir \rightarrow js} Z_{ir \rightarrow js} + \sum_s q_{ir \rightarrow s} Y_{ir \rightarrow s}, \forall ir. \quad (C3)$$

The adaptable production  $p_{ir}X_{ir}$  of a node  $ir$  has to equal the sum of its outgoing adaptable intermediate flows  $q_{ir \rightarrow js}Z_{ir \rightarrow js}$  plus the sum of its outgoing adaptable final demand flows  $q_{ir \rightarrow s}Y_{ir \rightarrow s}$ .

This constraint has to be fulfilled by each node, therefore the number of production output balance constraints amounts to  $n_{C3} = n_{rs} = n_r \cdot n_s$ .

### Exogenous forcing constraint

One specific regional sector's production is disrupted:

$$p_{i^*r^*} = \hat{p}. \quad (C4)$$

This constraint represents the effect of an exogenous forcing that affects one single node  $i^*r^*$  and sets its production ratio to a fixed value  $\hat{p} = (1 - \epsilon)$  with  $0 \leq \hat{p} \leq 1$ . Since this constraint only exists for one node,  $n_{C4} = 1$  holds.

### Number of constraints and free variables

If we consider a fully connected network, in which each regional sector is connected with all other regional sectors and all final demand nodes, the maximal amount of free variables  $n$  consists of the double amount of production ratio variables<sup>2</sup> plus the maximal possible amount of intermediate flows  $n_Z = n_{rs}^2$  plus the maximal possible amount of final demand flows  $n_Y = n_r \cdot n_{rs} = n_r^2 \cdot n_s$ :

$$n = n_Z + n_Y + 2 n_{rs} \quad (4.8)$$

$$= n_{rs}^2 + n_r^2 \cdot n_s + 2 n_{rs}. \quad (4.9)$$

The maximal amount of constraints  $m$  accounts to

$$m = n_{C1} + n_{C2} + n_{C3} + n_{C4} \quad (4.10)$$

$$= n_s^2 \cdot n_r + 2 n_{rs} + 1. \quad (4.11)$$

---

<sup>2</sup>The dublication is needed to transform the absolute value target function into a linear function, see Chap. 2.2.3

### 4.1.3 Special linear problem (LPS)

---


$$\min_{p_{ir}} F = \sum_{ir} |p_{ir} - 1| \quad (\text{T1})$$

$$s.t. \quad \sum_r q_{i^*r \rightarrow s} Y_{i^*r \rightarrow s} = \sum_r Y_{i^*r \rightarrow s} ; \forall s \quad (\text{C1}^*)$$

$$\sum_r q_{i^*r \rightarrow js} Z_{i^*r \rightarrow js} = p_{js} \sum_r Z_{i^*r \rightarrow js} ; \forall js \quad (\text{C2}^*)$$

$$p_{i^*r} X_{i^*r} = \sum_{js} q_{i^*r \rightarrow js} Z_{i^*r \rightarrow js} + \sum_s q_{i^*r \rightarrow s} Y_{i^*r \rightarrow s} ; \forall r \quad (\text{C3a}^*)$$

$$p_{ir} X_{ir} = \sum_{js} p_{js} Z_{ir \rightarrow js} + \sum_s Y_{ir \rightarrow s} ; \forall i \neq i^*; \forall r \quad (\text{C3b}^*)$$

$$p_{i^*r^*} = \hat{p}. \quad (\text{C4}^*)$$


---

#### Difference to the general linear program

Reducing the amount of free variables  $\mathbf{x}$  of the optimisation problem by transforming some of them into leading variables (which are dependent on the free variables but not part of the optimisation procedure), one can formulate new rules of assumed economical behaviour of the resulting network and save computing time.

Here, the focus of adaptation is set on flows transporting goods  $i^*$ , which shall remain freely adaptable and therefore stay part of the free variables  $\mathbf{x}$

$$\mathbf{x} = (\mathbf{p}, (q_{i^*r \rightarrow js}), (q_{i^*r \rightarrow s})) , \forall r, \forall js, \forall s, \quad (4.12)$$

while all other flows with  $i \neq i^*$  are considered as leading variables that are determined by predefined equations

$$q_{ir \rightarrow s} = 1, \forall i \neq i^*, \forall r, \forall s, \quad (4.13)$$

$$q_{ir \rightarrow js} = p_{js}, \forall i \neq i^*, \forall r, \forall js. \quad (4.14)$$

This formulation of the linear problem keeps all final demand flows  $Y_{ir \rightarrow s}$  with goods  $i \neq i^*$  constant. In addition, all intermediate input flows  $Z_{ir \rightarrow js}$  with  $i \neq i^*$  are assumed to be equal to the production ratio  $p_{js}$  of the target node  $js$ . This means that a change of production  $p_{js}$  and therefore a change of inputs  $Z_{ir \rightarrow js}$  is evenly distributed over all suppliers of node  $js$  and not adjusted through the optimisation

## 4 Model setup

procedure, as it is the case in the general linear problem.

### Final demand constraint

$$\sum_r q_{i^*r \rightarrow s} Y_{i^*r \rightarrow s} = \sum_r Y_{i^*r \rightarrow s}, \forall s. \quad (\text{C1}^*)$$

The final demand constraint (C1) only needs to be formulated for flows transporting good  $i^*$ , since the predefined equations (4.13) determine all final demand flows with  $i \neq i^*$ , making the corresponding constraints of the LPG dispensable. The maximal amount of final demand constraints reduces to  $n_{\text{C1}^*} = n_r$ .

### Supply scaling constraint

$$\sum_r q_{i^*r \rightarrow js} Z_{i^*r \rightarrow js} = p_{js} \sum_r Z_{i^*r \rightarrow js}, \forall js. \quad (\text{C2}^*)$$

Similarly, the supply scaling constraint (C2) is reduced to those equations using flows with good  $i^*$ . All other intermediate flows are determined by Eq. (4.14), leading to omission of the corresponding equations of the LPG and to a maximal amount of supply scaling constraints of  $n_{\text{C2}^*} = n_{rs} = n_r \cdot n_s$ .

### Production output balance constraint

$$p_{i^*r} X_{i^*r} = \sum_{js} q_{i^*r \rightarrow js} Z_{i^*r \rightarrow js} + \sum_s q_{i^*r \rightarrow s} Y_{i^*r \rightarrow s}, \forall r, \quad (\text{C3a}^*)$$

$$p_{ir} X_{ir} = \sum_{js} p_{js} Z_{ir \rightarrow js} + \sum_s Y_{ir \rightarrow s}, \forall i \neq i^*, \forall r. \quad (\text{C3b}^*)$$

The production output balance constraint (C3) is split into two constraints (C3a<sup>\*</sup>) and (C3b<sup>\*</sup>). The first is valid for nodes which produce the good  $i^*$  and keeps the form of the general linear problem with adaptability of intermediate flows as well as final demand flows. Contrarily, for nodes with  $i \neq i^*$ , constraint (C3b<sup>\*</sup>) exhibits reduced adaptability, since both flow variables are substituted by the predefined equations (4.13) (4.14), only leaving the production ratios  $\mathbf{p}$  as free variables.

There is exactly one production output balance constraint for each node  $n_{\text{C3}^*} = n_{rs}$ .

### Exogenous forcing constraint

$$p_{i^*r^*} = \hat{p}. \quad (\text{C4}^*)$$

This constraint remains identical to Eq. (C4) as in the LPG with  $n_{\text{C4}^*} = 1$ .

### Number of constraints and free variables

Altogether, the maximal amount of constraints for the LPS adds up to

$$m = n_{C1} + n_{C2} + n_{C3} + n_{C4} \quad (4.15)$$

$$= n_r + 2 n_{rs} + 1. \quad (4.16)$$

The maximal amount of intermediate flow variables is reduced to  $n_Z = n_{rs} \cdot n_r = n_r^2 \cdot n_s$ , since each node gets at maximum one input from each regional sector  $i^*r$ . Moreover, the amount of free final demand flow variables equals  $n_Y = n_r^2$ . Adding the doubled amount of production ratios, the maximal amount of variables for LPS is

$$n = n_Z + n_Y + n_{rs} \quad (4.17)$$

$$= n_r^2 \cdot n_s + n_r^2 + 2 n_{rs} . \quad (4.18)$$

## 4.2 Preprocessing of EORA data

The EORA data set is preprocessed to match model input data requirements.

### Rounding

All entries of the EORA26 MRIOT  $\mathbf{M}^a = (\mathbf{Z}, \mathbf{Y})$  are rounded to the third decimal:

$$M_{xy}^b := \frac{\text{round}(M_{xy}^a \cdot 10^3)}{10^3}, \forall x, y, \quad (4.19)$$

where  $M_{xy}$  is the matrix entry in row  $x$  and column  $y$ .

### Threshold

Only entries with a value higher than the threshold  $\tau$  are taken into account. Due to harmonization effects of the MRIOT creation process (Lenzen et al., 2013), the trade flows are likely to contain a harmonization error that could distort small flow values significantly. For this reason, we only consider trade flows above the threshold

## 4 Model setup

$\tau$  as relevant and sufficiently accurate. In this thesis a threshold of  $\tau = 10^6$  \$/year is used:

$$M_{xy}^c := \begin{cases} M_{xy}^b & M_{xy}^b > \tau \\ 0 & M_{xy}^b \leq \tau \end{cases}, \forall x, y. \quad (4.20)$$

### Nodes with inappropriate inputs

If a regional sector has no input  $D_{ir} = 0$  or the total input exceeds its total output  $D_{ir} > X_{ir}$ , all of its input and output flows are set to zero. This ensures that the trade flows can be interpreted as elements of supply chains with regional sectors creating a value added. The removed flows are assumed to be influenced by harmonization effects or to occur as a result of particular political frameworks, like the existence of high subsidies for instance.

$$M_{xy}^d := \begin{cases} 0 & D_x = 0 \vee D_x > X_x \\ 0 & D_y = 0 \vee D_y > X_y \\ M_{xy}^c & \text{otherwise} \end{cases}, \forall x, y. \quad (4.21)$$

### Result data set

The resulting data set after the application of the above preprocessing methods shows a much smaller matrix density  $\rho_Z$  and  $\rho_Y$  than the original data set (see Fig. 4.1 that depicts the heat map of the  $Z$  matrix of the year 2011). An overview of the densities of the  $Z$  and  $Y$  matrices and the resulting linear problem dimensions for each sample year is given in Tab. 4.1.

year	2011	2006	2001	1996
$\rho_Z$	0.0214	0.0170	0.0109	0.0102
$\rho_Y$	0.0930	0.0726	0.0517	0.0507
$n^{\text{LPG}}$	$\approx 6 \cdot 10^6$	$\approx 5 \cdot 10^6$	$\approx 5 \cdot 10^6$	$\approx 3 \cdot 10^6$
$m^{\text{LPG}}$	$\approx 8 \cdot 10^5$	$\approx 8 \cdot 10^5$	$\approx 8 \cdot 10^5$	$\approx 7 \cdot 10^5$
$n^{\text{LPS}}$	$\approx 3 \cdot 10^5$	$\approx 6 \cdot 10^5$	$\approx 6 \cdot 10^5$	$\approx 4 \cdot 10^5$
$m^{\text{LPS}}$	$\approx 7 \cdot 10^4$	$\approx 8 \cdot 10^4$	$\approx 8 \cdot 10^4$	$\approx 8 \cdot 10^4$

Table 4.1: Matrix densities and linear problem dimensions of LPG and LPS for all four sample years.

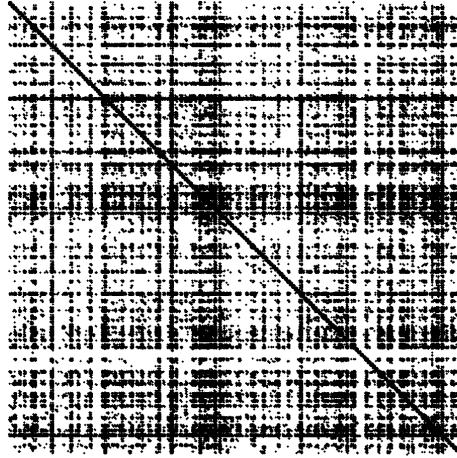


Figure 4.1: Heat map of the preprocessed intermediate flow matrix  $\mathbf{Z}$  of EORA26 of the year 2011. An entry represents a flow between two regional sectors. While the original unprocessed EORA26 matrix contains flow values in each cell ( $\rho_Z = 1$ ), the preprocessed one comes along with a significantly lower matrix density. The diagonal represents domestic trade flows, which are practically present for every region, while the other entries represent strong foreign trade flows.

### 4.3 Aggregated version of EORA data

One simulation run of the optimisation problem LPG using the preprocessed EORA26 data  $\mathbf{M}^d$  takes about two days of computing time<sup>3</sup>. Since we would analyse statistical properties of the network by affecting each regional sector once with the exogenous forcing and thus need to conduct a great number of simulation runs, a reduction of complexity is achieved by aggregating some of the regions. We keep 28 selected regions with a high GDP value or interesting geographical locations as independent regions, while the other regions are aggregated into four new proxy regions, namely *rest of Europe*, *rest of America*, *rest of Asia* and *rest of Africa*. Table A.3 lists all the regions of the new data set  $\mathbf{M}^{\text{agg}}$ .

By reducing the amount of regions from 186 to 32, we also reduce the amount of constraints  $m$  and variables  $n$ . Since the runtime of the simplex algorithm is practically controlled by  $\mathcal{O}(m)$ , we are able to reduce it to about 1.5 hours for LPG and 30 min for LPS, what allows for statistical analysis. The densities and linear problem dimensions of the aggregated EORA26 data set  $\mathbf{M}^{\text{agg}}$  are shown in Tab. 4.2.

---

<sup>3</sup>using the high performance cluster computers of PIK, see Sec. 4.4 for cluster specifications

#### 4 Model setup

	2011	2006	2001	1996
$\rho_Z$	0.2517	0.2212	0.1599	0.1512
$\rho_Y$	0.5811	0.5382	0.4697	0.4615
n (LPG)	$\approx 2 \cdot 10^6$	$\approx 2 \cdot 10^6$	$\approx 1 \cdot 10^6$	$\approx 1 \cdot 10^6$
m (LPG)	$\approx 2 \cdot 10^5$			
n (LPS)	$\approx 2 \cdot 10^5$	$\approx 6 \cdot 10^5$	$\approx 6 \cdot 10^5$	$\approx 4 \cdot 10^5$
m (LPS)	$\approx 2 \cdot 10^4$	$\approx 8 \cdot 10^4$	$\approx 8 \cdot 10^4$	$\approx 8 \cdot 10^4$

Table 4.2: Densities of aggregated EORA26 data and dimensions of the corresponding linear problems.

#### 4.4 Technical implementation

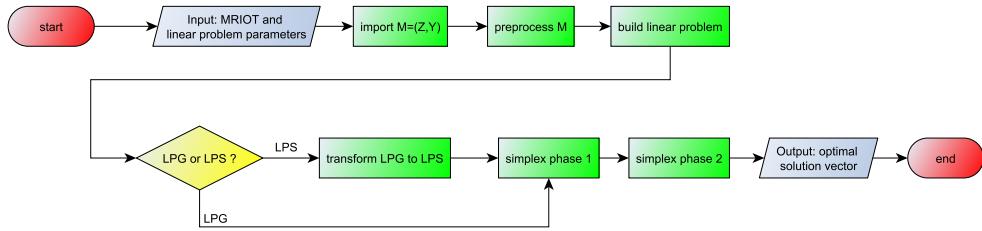


Figure 4.2: Flowchart of main steps to set up the linear problem and to conduct the optimisation procedure.

The code is implemented in MATLAB<sup>TM</sup> (MATLAB, 2012) and covers six main functional steps (see Fig 4.2). After starting, the software imports the MRIOT data as well as scenario variables (e.g.  $i^*$ ,  $r^*$ , LPS or LPG). In the next step, the data is preprocessed and prepared to be translated into the general linear optimisation problem (LPG), setting the vector of free variables  $\mathbf{x}$ , the constraint matrix  $\mathbf{A}$ , the constraint vector  $\mathbf{b}$  and the cost vector  $\mathbf{c}$ . If the scenario definitions demand the special linear problem LPS instead of the general one, a routine for transforming the LPG to LPS is called. Finally, the simplex algorithm is applied to the linear problem, starting with phase I to determine an initial basic feasible solution and afterwards finding the optimal solution by using the simplex phase II method. As a result, either the optimal solution vector  $\bar{\mathbf{x}}^*$  or a message about infeasibility of the optimisation problem is saved in the output file. The software runs were conducted on high performance cluster computers<sup>4</sup> at PIK.

<sup>4</sup>with 3.4 GHz and 4 GByte DDR4 memory per core.

# 5 Application of the linear problem to a stylised network

## 5.1 Exemplary optimisation results

The linear problem defined in the section above is applied to a stylized sample MRIOT and its corresponding network (see Fig. 5.1) with three regions and two sectors. The amount of nodes was chosen to be as small as possible to get a comprehension of the effects the constraints and the target function have on the optimisation results. The regional sector  $i^*r^* = 11$  is affected by a forcing of  $\epsilon = 0.5$ . The simplex method is used to determine the network response once for the general problem LPG (see Fig. 5.2) and once for the special problem LPS (see Fig. 5.3).

Since the LPG allows more adaptivity, the adaptation effort  $F^{\text{LPG}} = 1.125$  is lower than the result for the LPS case  $F^{\text{LPS}} = 1.25$ .

Both variant results continue satisfying the constant final demand, fulfilling constraint (C1). Moreover, the *Leontief production function* (C2) is satisfied in both cases, since a reduced (augmented) production leads to a proportional decrease (increase) in inputs, with constant shares of input flows in the total input. Additionally, the production output balance constraint (C3) is satisfied, since the sum of a node's outputs equals its new total production  $X^1 = pX^0$ .

The resulting flows differ between LPG and LPS. In the latter, the final demand flows of goods  $i = 2 \neq i^*$  stay constant and inputs of nodes producing these goods are not freely adaptable, but change with the rate of their source node's production change  $p$ , as stated by the equations (4.13) and (4.14).

By taking a look at the results for the LPG, one observes that the output flows of regional sector  $ir = 21$  are rearranged due to the reduced demand of  $i^*r^* = 11$ . Furthermore, regional sector  $ir = 23$  does not receive enough inputs of good  $i = 1$  anymore to keep its production constant ( $p_{23} = 0.75$ ). The lack of final demand flows of good  $i^* = 1$  into region  $r1$  is compensated by  $ir = 12$ , that exhibits the strongest positive production change with  $p_{12} = 1.375$ .

## 5 Application of the linear problem to a stylised network

region		r1	r1	r2	r2	r3	r3	r1	r2	r3	Total output
sector	index	s1	s2	s1	s2	s1	s2	FD1	FD2	FD3	
	ir=11	ir=21	ir=12	ir=22	ir=13	ir=23	r1	r2	r3		
r1	s1	10					10	10		30->15	
	s2	10	10					10			30
r2	s1			10				10	10		30
	s2				10			10	10	10	40
r3	s1					10				10	20
	s2						10			20	30
Input s1		10		10		10	10	20	10	10	
Input s2		10	10	10		10	10	20	10	30	
Total input		20	10	10	10	10	20	40	20	40	

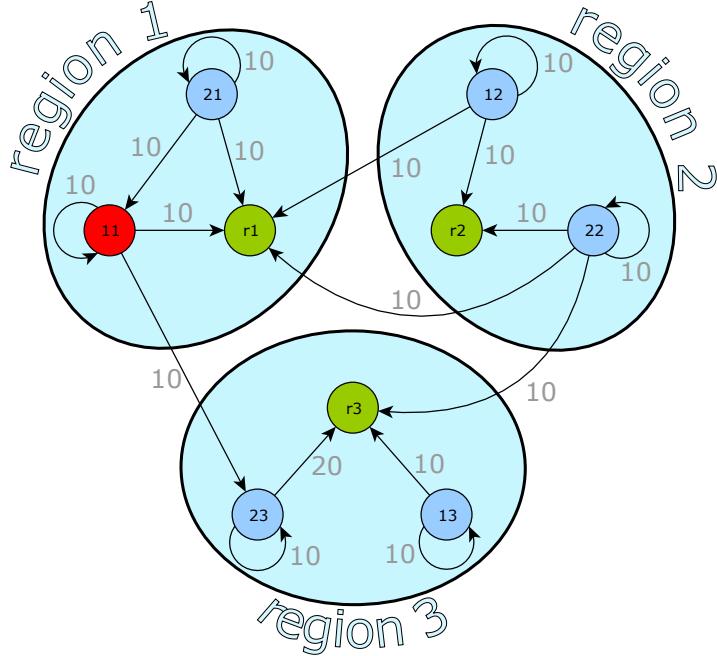


Figure 5.1: Sample MRIOT and corresponding network with three regions and two sectors before the adaptation process. The amount of nodes was kept as low as possible to give a good comprehension of the effects the constraints and target function have on the optimisation results. The regional sector  $i^*r^* = 11$  (red) is subject to a forcing  $\epsilon = 0.5$ , reducing its output by half from  $X_{i^*r^*}^0 = 30$  to  $X_{i^*r^*}^1 = 15$ .

The resulting production ratios of the optimisation procedure for the LPG are  $\mathbf{p} = (p_{11}, p_{21}, p_{12}, p_{22}, p_{13}, p_{23}) = (0.5, 1, 1.375, 1, 1, 0.75)$ . Since the network consists of six regional sectors and 17 flows, the general linear problem contains  $n = 29$  free variables (with the production ratios  $p$  counted twice due to the absolute value target function). The amount of constraints is  $m = 21$ , consisting of six final demand constraints, eight supply scaling constraints, six production output balance constraints and one exogenous forcing constraint.

### 5.1 Exemplary optimisation results

For the LPS, the resulting production ratios are  $\mathbf{p} = (p_{11}, p_{21}, p_{12}, p_{22}, p_{13}, p_{23}) = (0.5, 0.75, 1.5, 1, 1, 1)$ . The amount of free variables in the LPS is smaller than in the LPG. Only flows with  $i = i^*$  remain freely adaptable, resulting in  $n = 8$ . The amount of constraints is  $m = 14$ , three final demand constraints, four supply scaling constraints, six production output balance constraints and one exogenous forcing constraint.

region			r1	r1	r2	r2	r3	r3	r1	r2	r3	Total output
	sector		s1	s2	s1	s2	s1	s2	FD1	FD2	FD3	
		index	ir=11	ir=21	ir=12	ir=22	ir=13	ir=23	r1	r2	r3	
r1	s1	ir=11	5				7.5	2.5				15
	s2	ir=21	5	10				15				30
r2	s1	ir=12			13.75			17.5	10			41.25
	s2	ir=22			10			5	10	15		40
r3	s1	ir=13				10			10			20
	s2	ir=23					7.5			15		22.5
Input s1		5		13.75		10	7.5	20	10	10		
Input s2		5	10		10		7.5	20	10	30		
Total input		10	10	13.75	10	10	15	40	20	40		

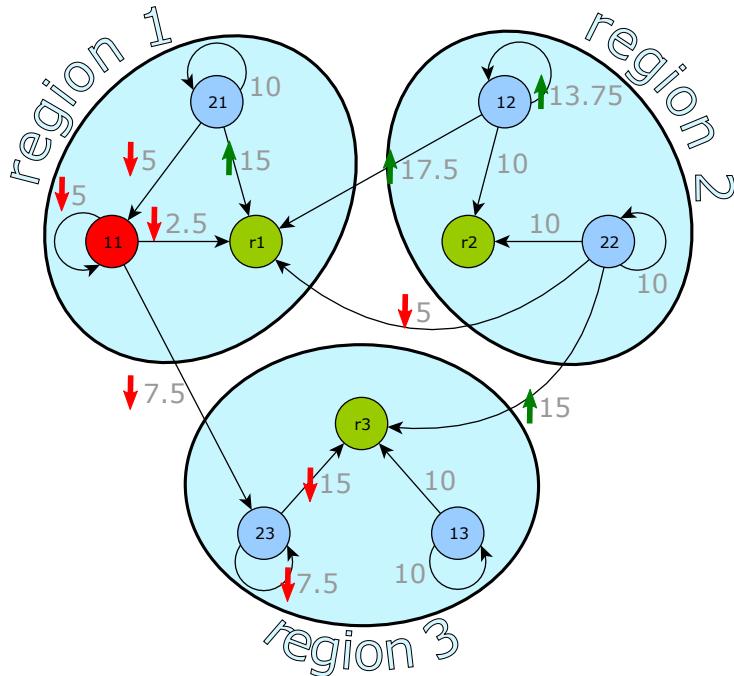


Figure 5.2: Adapted MRIOT and network according to the general linear problem LPG. The regional sector  $ir = 12$  is increasing its production ( $p_{12} = 1.375$ ) to compensate the production loss of  $ir = 11$ . Regional sector  $ir = 21$  is rearranging its flows, since  $i^*r^* = 11$  only needs half of its original inputs  $Z_{21 \rightarrow 11}$ . Furthermore,  $ir = 23$  receives less inputs from  $i^*r^* = 11$  and has to reduce its production as a consequence, what leads to an increased flow from regional sector  $ir = 22$  to region  $r3$  to keep the final demand inputs of good  $i = 2$  constant.

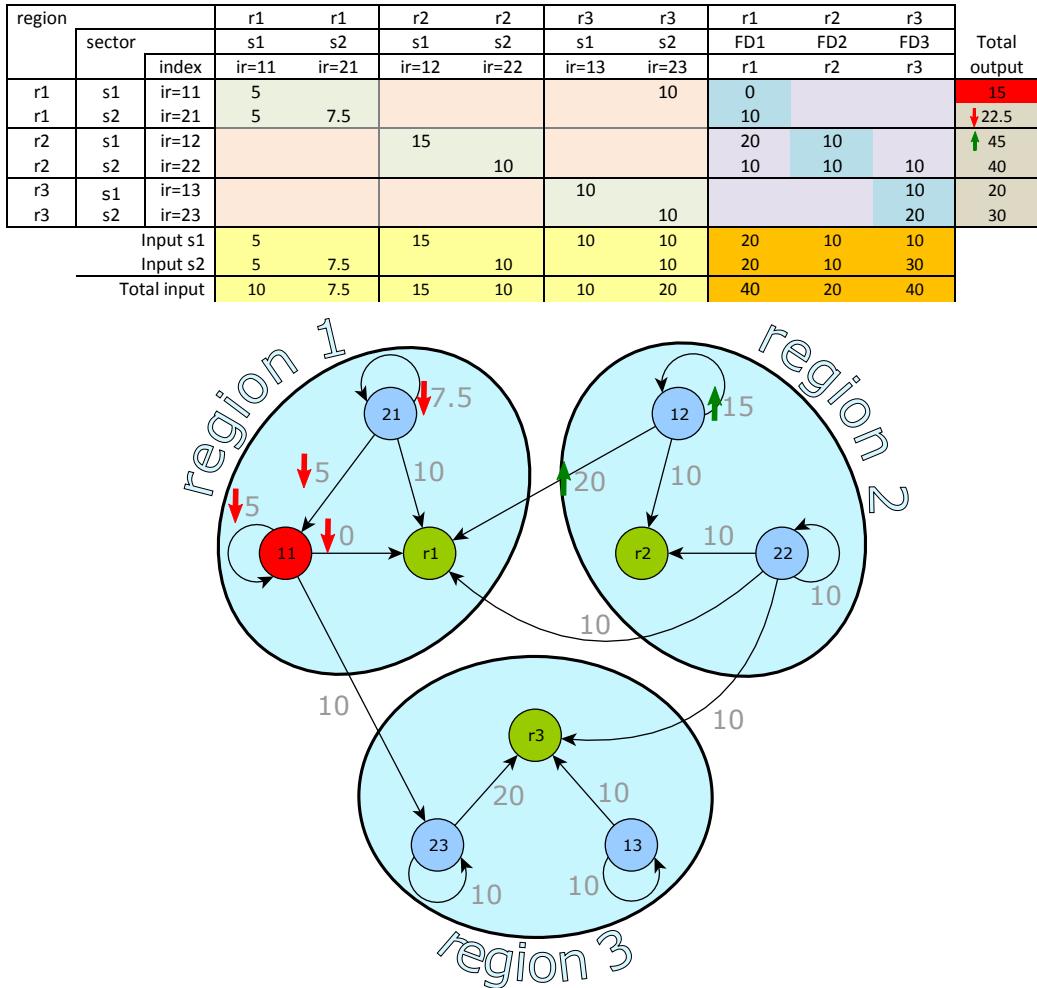


Figure 5.3: Adapted MRIOT and network according to the special linear problem LPS. Regional sector  $ir = 21$  is forced to reduce its inputs to  $i^*r^* = 11$  by half, since all intermediate input flows with  $i \neq i^*$  change with the target's production ratio. All final demands with  $i \neq i^*$  are fixed, so  $ir = 21$  needs to reduce its production  $p_{21} = 0.75$ . Regional sector  $ir = 12$  compensates the lack of final demand flows into region  $r1$  while still providing goods for  $r2$  and therefore increasing its production  $p_{12} = 1.5$ .

# 6 Optimisation results

## 6.1 Adaptation effort for different failure strengths

The exogenous forcing  $\epsilon$  disrupts the network, which has to expend adaptation effort  $F(\epsilon)$  (see definition in Chap. 4.1.1) to continue fulfilling all constraints. Simulation runs of various values of  $\epsilon$  for a given set of framework parameters (year,  $i^*$ ,  $r^*$ , LPG) are used to analyse  $F(\epsilon)$  (see Fig. 6.1). For small perturbations  $\epsilon$ , the function is strictly linear, which is an expected behaviour for a linear optimisation problem, since a variation of  $\epsilon$  only changes constraint (C4) linearly. In this case, the hyperplane corresponding to constraint (C4) is moved parallel to the former hyperplane position. The optimal result  $\bar{x}^*$  lies on a corner of the polyhedron and moving the hyperplane (C4) moves the position of this corner, while not altering the topology of the polyhedron. For that reason, small variations of  $\epsilon$  do not necessarily exchange the optimal corner, but only scale the components of  $\bar{x}^*$  linearly.

It is possible, though, to leave the linear domain of  $F(\epsilon)$  by increasing  $\epsilon$  until a critical point  $\epsilon_{c,i^*r^*}$  is traversed. In this case, the former optimal corner solution of the polyhedron is moved so far, that another corner gets more cost effective. This implies a new set of active constraints. The new corner in general consists of components without any linear relation to the former corner solution. It is possible that the frequent change of the optimal corner solution after crossing  $\epsilon_{c,i^*r^*}$  is related to numerical errors of the simplex algorithm. The domain's significance should not be over-interpreted until further investigation of the numerical errors is conducted.

After reaching the critical value  $\epsilon_{c,i^*r^*}$ , that is specific for each affected regional sector  $i^*r^*$ , the target function increases nonlinearly until it becomes infeasible after a value of  $\tilde{\epsilon}_{i^*r^*}$ . The value of  $\tilde{\epsilon}_{i^*r^*}$  seems to represent an essential amount of production  $X_{i^*r^*}$  that is necessary to maintain feasibility of the system. Especially regional sectors of very low values  $\tilde{\epsilon}_{i^*r^*}$  are of interest, since these lead to an infeasible solution even for relatively slight disruptions.

A similar behaviour is found when investigating only one summand  $p_{113}(\epsilon)$  of the target function (see Fig. 6.1, green curve). Here, we also see a linear behaviour,

## 6 Optimisation results

until  $\epsilon_{c,i^*r^*}$  is reached. Crossing it,  $p_{113}(\epsilon)$  exhibits linearity in intervals of  $\epsilon$  with different slopes for each of these intervals, until  $\epsilon$  traverses a final critical value and stays at  $p_{113} = 0$ , until the infeasible domain is reached at  $\tilde{\epsilon}_{i^*r^*}$ .

This qualitative behaviour - namely the existence of a linear, nonlinear and infeasible domain - can be observed for most of the other possibly affected regional sectors  $i^*r^*$ , of which four are presented in Fig. 6.2a.

We take a look at the linear domains of three different regional sectors for a given optimisation run with  $i^*r^*$  (see Fig. 6.2b) and observe that the regional sectors change their production ratio  $p_{ir}(\epsilon)$  in the linear domain with different slopes. The slope describes the sensitivity of the unaffected regional sectors to the forcing  $\epsilon$ .

### Infeasibility limit $\tilde{\epsilon}_{i^*r^*}$

Not all optimisation problem formulations are always feasible or exhibit an optimal solution. Apart from the possibility of being unbounded, it could happen that no solution for the problem exists since the constraints can not be satisfied under the given parameters. Especially the value of the forcing  $\epsilon$  plays an important role, since we observe that a choice of  $\epsilon > \tilde{\epsilon}_{i^*r^*}$  leads to an infeasible optimisation problem. An overview about the value distribution of  $\tilde{\epsilon}_{i^*r^*}$  for a parameter set of the LPG for the year 2011 is given in Figure 6.3.

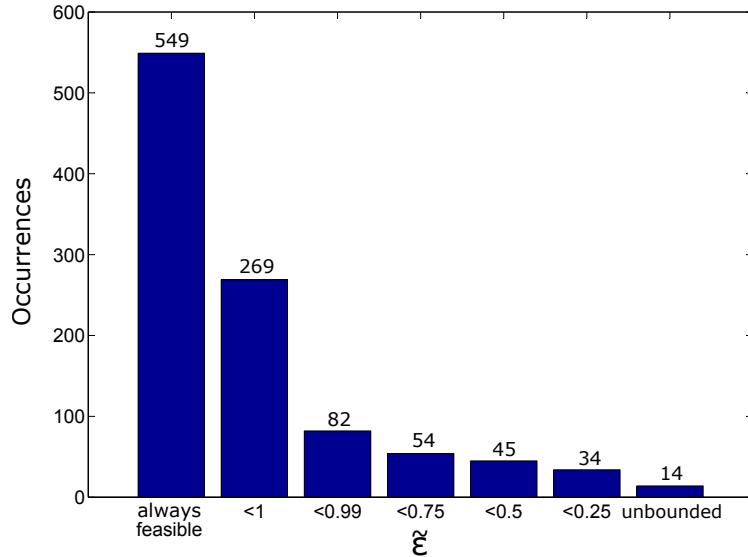


Figure 6.3: Cumulative distribution of  $\{\tilde{\epsilon}_{i^*r^*}\}$  for the LPG in the year 2011. Most of the regional sectors (549) lead to a solvable optimisation problem. It is unbounded in 14 cases.

### 6.1 Adaptation effort for different failure strengths

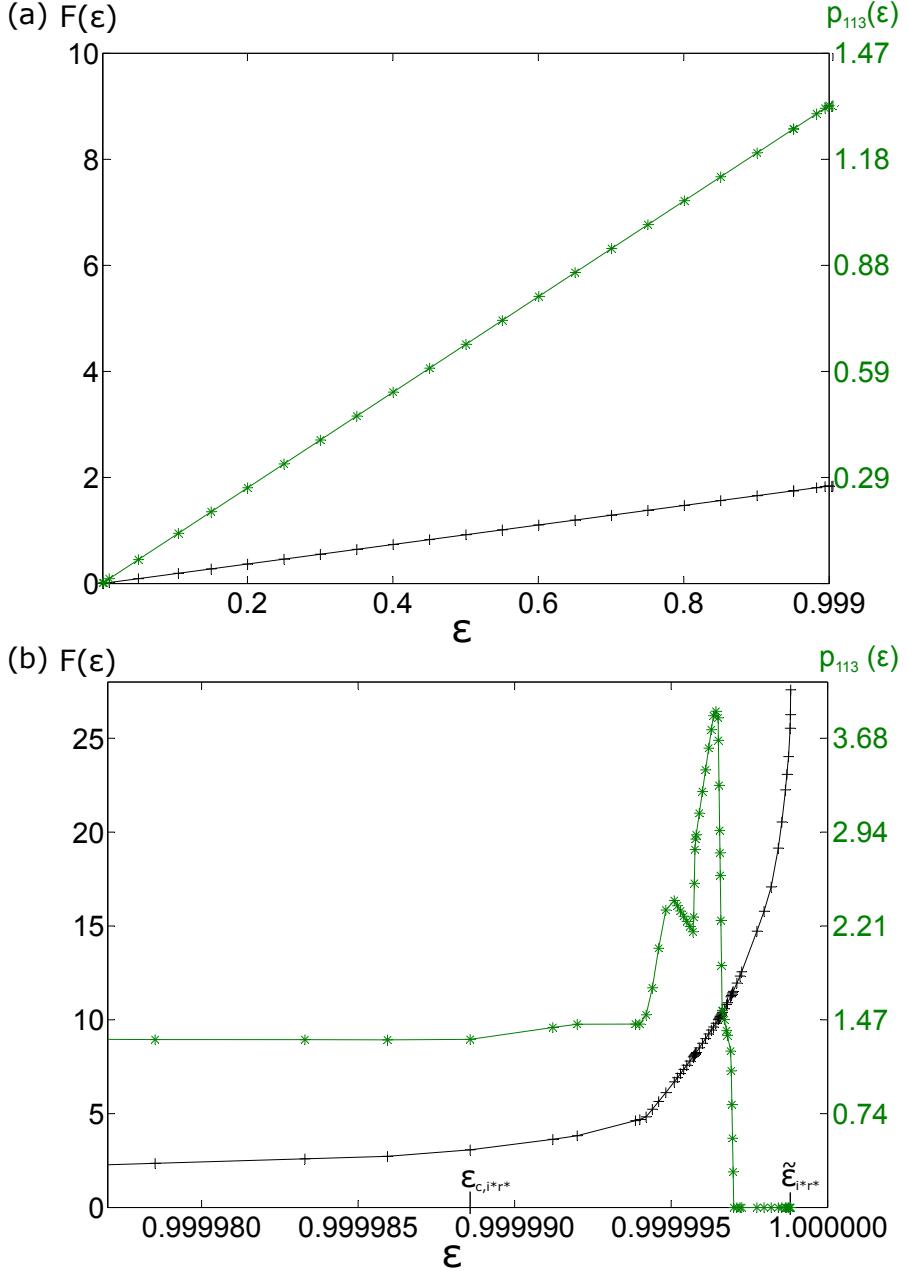


Figure 6.1: (a) Linear domain of  $F(\epsilon)$  and  $p_{113}(\epsilon)$ , using the LPG for the year 2011. The affected regional sector  $i^*r^*$  is Japan Machinery. The response of the regional sector  $ir = 113$  (China Machinery) is shown representatively for the behaviour of a regional sector's production ratio  $p(\epsilon)$ . (b) Linear, nonlinear and infeasible domains of both functions. For values  $\epsilon_{c,i^*r^*} < \epsilon < \tilde{\epsilon}_{i^*r^*}$  the production ratio  $p_{113}(\epsilon)$  is linear in intervals of  $\epsilon$  with specific slopes for each of these intervals.

## 6 Optimisation results

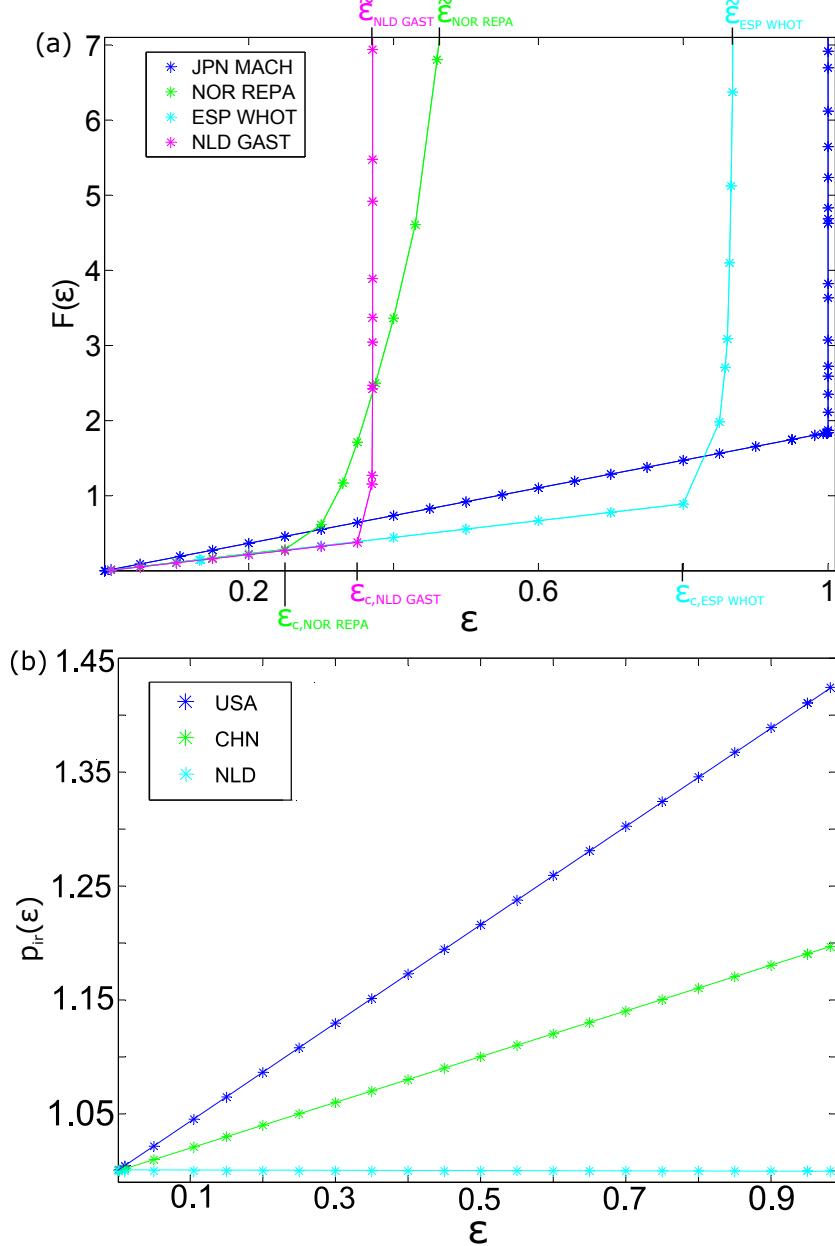


Figure 6.2: (a) Target function  $F(\epsilon)$  for several simulation runs with different  $i^*r^*$ . For each of the four runs, the linear, nonlinear and infeasible domains are easy to identify. To each  $i^*r^*$  belongs a specific  $\epsilon_{c,i^*r^*}$  and  $\tilde{\epsilon}_{i^*r^*}$ . (b) Linear domain of the target functions of different regional sectors  $p_{ir}(\epsilon)$  for a simulation run with JPN MACH as affected regional sector  $i^*r^*$ . The MACH sectors of USA and China increase their production, while other regions, like NLD, keep it constant.

### 6.1 Adaptation effort for different failure strengths

Since it is unlikely that a disruption  $\epsilon$  affects nearly the whole production of a country, we take a closer look at regional sectors with  $\tilde{\epsilon}_{i^*r^*}$  in the range of  $[0, 0.5]$ . We observe that 18 of all 32 regions contain a regional sector within that range, of which especially Argentina, New Zealand, Italy, Spain and Sweden have more than three regional sectors therein (see Fig. 6.4 and Tab. 6.1). Among the sectors, the final demand oriented sectors CONS, REPA, GAST, ADMI, HOUS and the OTHE sector dominate. Since we assume the final demand to be constant and want the system to continue satisfying it, it might not be possible to find alternative suppliers of specific final demands in case of a disruption in the mentioned regional sectors with a high value of  $\tilde{\epsilon}_{i^*r^*}$ .

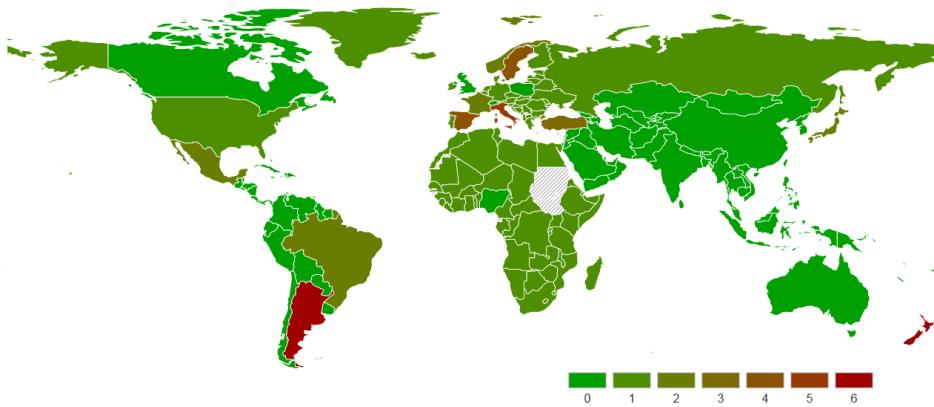


Figure 6.4: Occurrences of  $\tilde{\epsilon}_{i^*r^*} \leq 0.5$  for each region (LPG, year 2011).

Regions	ARG	NZL	ITA	ESP	SWE	TUR	BRA	FRA	JPN
Sectors	CONS	CONS	CONS	CONS	CONS	REPA	HOUS	CONS	CONS
REPA	REPA	REPA	REPA	REPA	REPA	HOUS	OTHE	ADMI	ADMI
GAST	GAST	GAST	GAST	GAST	ADMI	OTHE			
COMM	ADMI	ADMI	ADMI	ADMI	OTHE				
ADMI	HOUS	OTHE							
EDHE	OTHE								

Regions	MEX	NOR	DEU	NLD	RUS	ZAF	USA	REU	RAF
Sectors	REPA	ADMI	ADMI	GAST	REPA	OTHE	CONS	OTHE	HOUS
GAST	OTHE								

Table 6.1: Sectors of each region with  $\tilde{\epsilon}_{i^*r^*} \leq 0.5$  (LPG, year 2011).

## 6.2 Adaptation effort for different failure locations

In order to get an overview of the adaptation effort  $F(i^*r^*)$  the system expends, a total set of simulation runs is executed, where each existing regional sector  $ir$  is assumed to be affected once, while the other system parameters (LPG / LPS,  $\epsilon = 0.5$ , year 2011) stay constant. The aggregated EORA26 data set is used for this purpose with a total amount of  $n_{rs} = 832$  regional sectors. Only those simulation runs are taken into account where  $F(i^*r^*)$  is in the linear domain, making the results comparable. To determine the solution domains for the simulation set with  $\epsilon = 0.5$ , a second auxiliary set with  $\epsilon = 0.01$  is used. For a forcing of  $\epsilon = 0$ , no adaptation effort is expended since the network remains in the original state, so  $F(0) = 0$ . For a specific run with  $i^*r^*$ , the slope  $F'(\epsilon)$  is approximated by the difference quotient

$$F'(\epsilon) = \frac{F(0.01)}{0.01}. \quad (6.1)$$

To determine the domain of a simulation run with  $\epsilon = 0.5$ , the relative difference between the assumed linear result  $F' \cdot 0.5$  and  $F(0.5)$  is calculated. If it lies under a threshold of  $\tau = 5\%$ , the solution  $F(0.5)$  is assumed to be in the linear domain

$$\frac{F' \cdot 0.5}{F(0.5)} \begin{cases} \leq \tau = 0.05 & \text{linear domain} \\ > \tau = 0.05 & \text{nonlinear domain} \end{cases}. \quad (6.2)$$

The simulation runs within the nonlinear domain as well as infeasible runs are omitted in the figures presented in this chapter.

### Impacts using the LPG

Using the LPG, a forcing on a regional sector in China or the USA causes high network adaptation efforts  $F$  (see Fig. 6.5). Japan also shows a peak, though a much smaller one. Overview plots for the years 1996, 2001 and 2006 can be found in the appendix (Fig. A.4 - A.6). The adaptation effort to compensate for a failure in China increases with time while it decreases for failures of Japanese sectors. This is in line with the growing economic strength of China in the last 20 years.

## 6.2 Adaptation effort for different failure locations

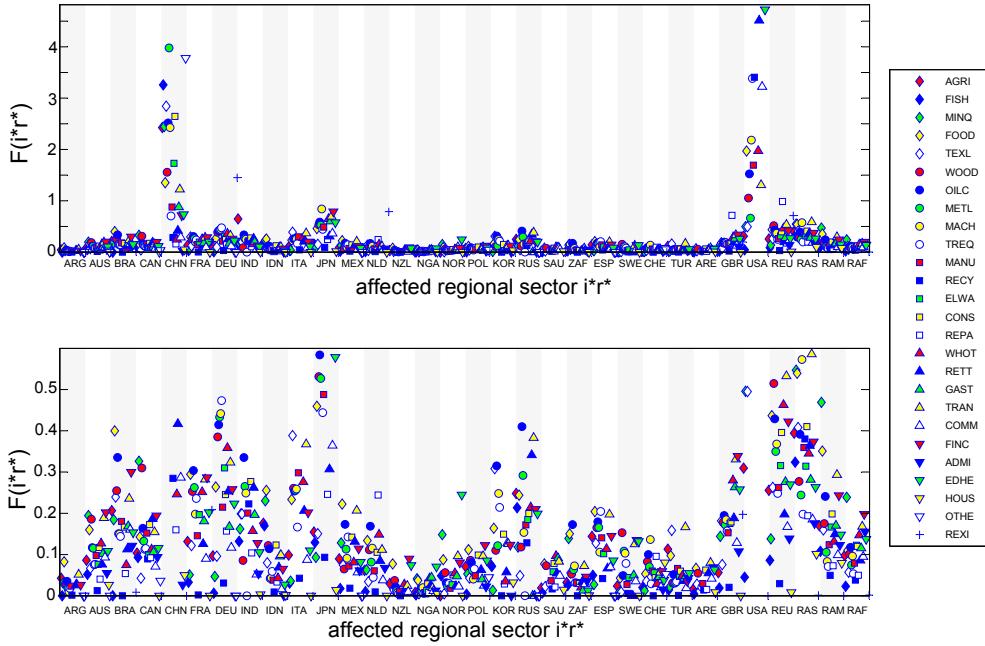


Figure 6.5: Network adaptation effort  $F(i^*r^*)$ , using the LPG, for all feasible simulation runs within the linear domain for the year 2011 and a forcing of  $\epsilon = 0.5$ . A failure of a Chinese or US sector induces a high network adaptation effort.

### Impacts using the LPS

The same figures were created for the LPS (Fig. 6.6). Since the flows are partly predetermined and adaptability is restricted to the flows of  $i^*$  goods, more regions cause a relatively high adaptation effort compared to the LPG results (not only China and the USA). The target function values for the LPS are always higher than for the LPG, since the latter allows for more adaptation with a higher degree of freedom than the LPS does (see Fig 6.7).

## 6 Optimisation results

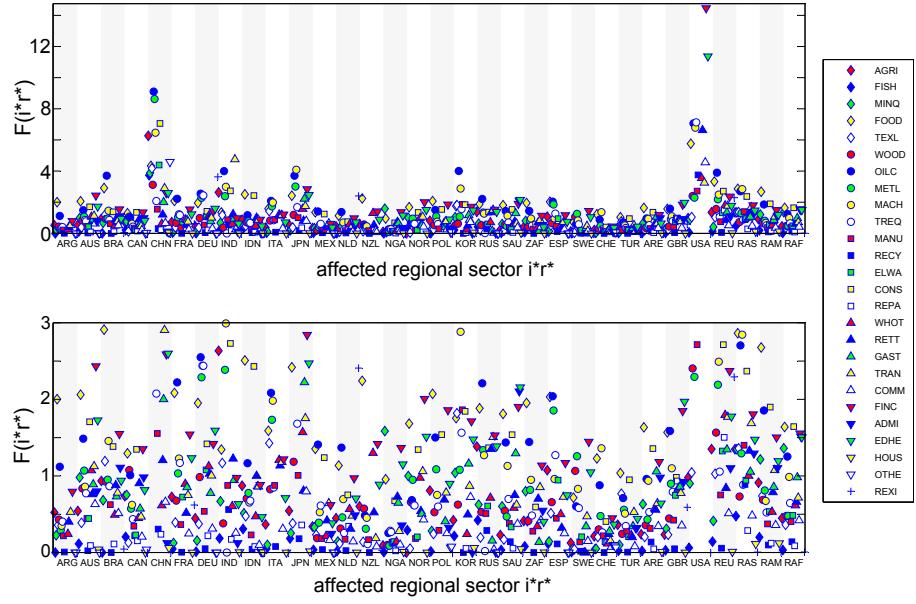


Figure 6.6: Results  $F(i^*r^*)$ , using the LPS, for all feasible simulation runs within the linear domain for the year 2011,  $\epsilon = 0.5$ . A failure of a Chinese or US sector induces a high network adaptation effort, but also Brasil, India, Japan and Korea have some regional sectors whose disruption leads to high values  $F(i^*r^*)$ .

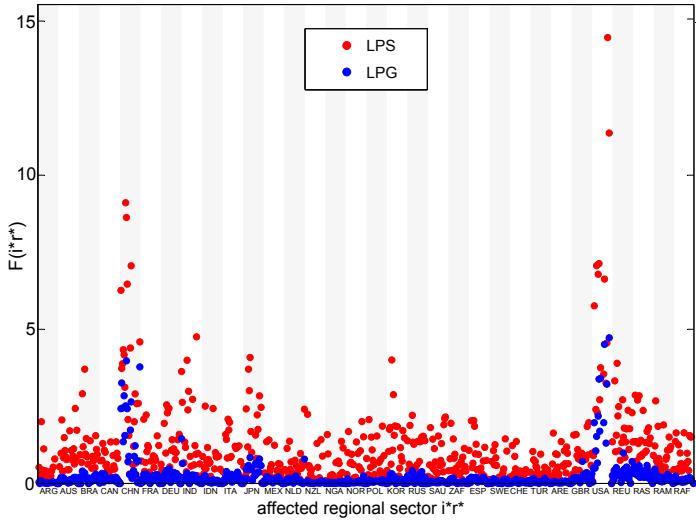


Figure 6.7: Comparison of network adaptation efforts of the LPS (red) and the LPG (blue): Simulation runs of the LPG always result in lower target function values than for the LPS, since the former has more degrees of freedom to find an optimal response to the network disruption.

### 6.3 Main compensators

Of all regional sectors within the affected sector  $i^*$ , the one with the strongest increase of  $p_{i^*r}$  is called the main compensator. The set of optimisation results for the year 2011, using the LPG and  $\epsilon = 0.5$ , is analysed and the main compensators for each simulation run are determined (Fig. 6.9). An overview with a combined regional and sectoral view for selected sectors is shown in Fig. 6.8. A pattern can be identified for the sectoral representation (Fig. 6.9b) with regional sectors from specific regions compensating nearly exclusively for specific sectors. For the regional representation (Fig. 6.9a) no such pattern can be detected.

For the majority of the problems studied, regional sectors from China or the USA act as the main compensators, while the Italian regional sector does so in the repair sector (REPA). Regional sectors from Japan (JPN) and rest of Asia (RAS) are main compensators in the case of USA failures. Moreover, Germany (DEU) compensates frequently in the re-export re-import sector (REXI).

More overviews of main compensators for the years 1996, 2001 and 2006 are

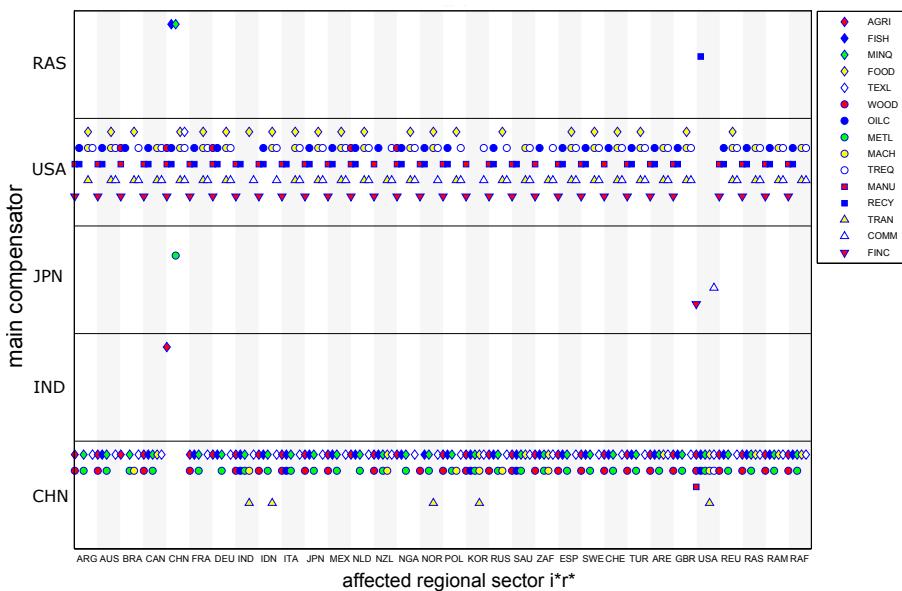


Figure 6.8: Combined regional and sectoral view of the main compensators for selected sectors and simulation runs of the year 2011,  $\epsilon = 0.5$ , using the LPG. Each feasible scenario with affected regional sector  $i^*r^*$  is displayed on the x axis and the corresponding main compensator on the y axis, while the markers signify the affected sector  $i^*$ . Regional sectors in China and the USA act as main compensators in most of the cases.

## 6 Optimisation results

presented in the appendix (Fig. A.7- A.9). The importance of China as main compensator increases with time, particularly in sectors such as mining (MINQ), textiles (TEXL) and metal products (METL).

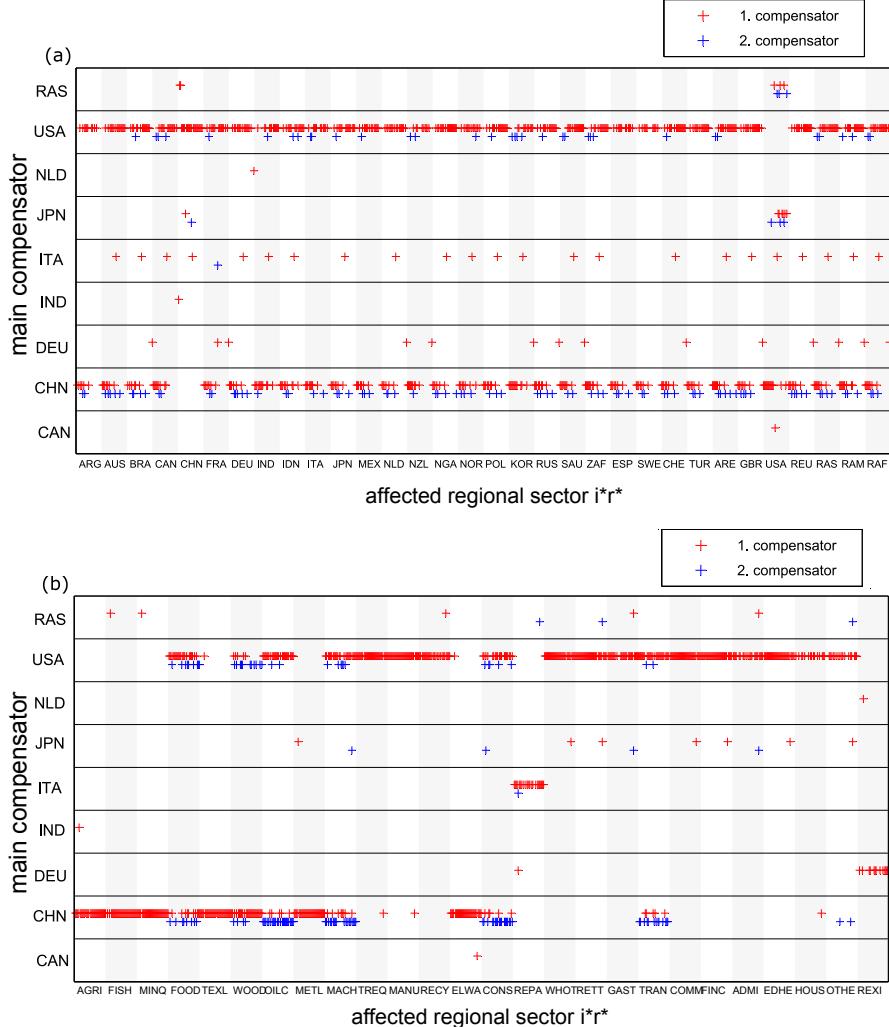


Figure 6.9: **(a)** Overview of the main compensators sorted by regions for the year 2011,  $\epsilon = 0.5$ , using the LPG and the aggregated EORA26 data, where each feasible scenario with affected regional sector  $i^*r^*$  is displayed on the x axis and the corresponding main compensator on the y axis. In most of the scenarios, the respective sector in China or the USA acts as main compensator, while some other regional sectors do so in specific single scenarios. **(b)** Same overview, sorted by sectors. Failures in specific sectors are nearly exclusively compensated by regional sectors from China (e.g. the AGRI sector) or the USA (e.g. the MANU sector).

### Sector assignment

The regional sectors of some countries compensate disruptions in specific sectors nearly exclusively, so we determine the share of these regions in the main compensators for all sectors (see Fig. 6.10).

The further advancing in time, the more sectors are compensated by the respective regional sector in China and less by the one in the USA.

We observe that in 1996 the regional sectors in China only compensate for few sectors (AGRI, FISH), extending its share in different sectors when advancing in time (MINQ, TEXL, METL) and partly acting as main compensator for sectors such as OILC and MACH in 2011.

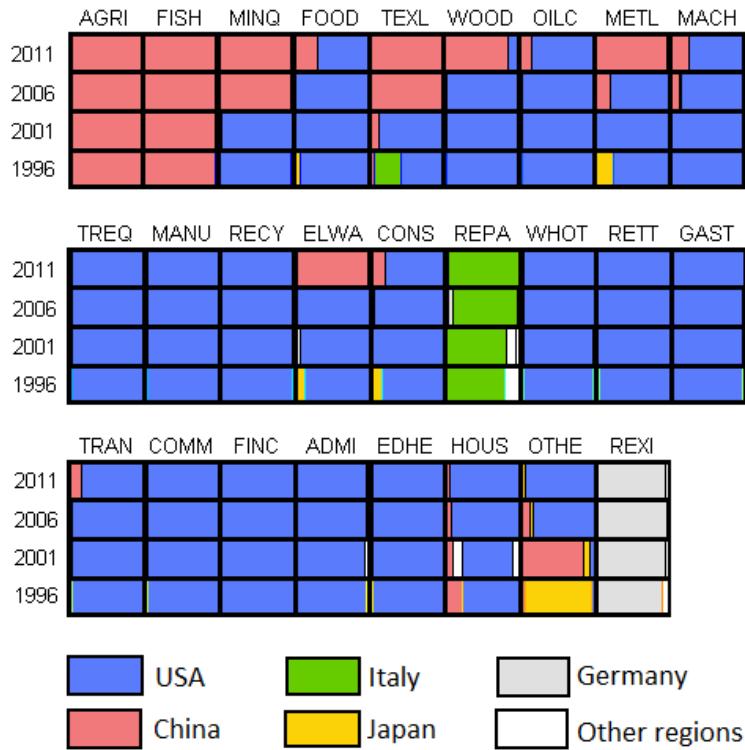


Figure 6.10: Share of regions in main compensators for each sector of the years 2011, 2006, 2001 and 1996. Results taken from simulation runs with parameters  $\epsilon = 0.5$  and using the LPG. In 1996 sectors in Japan and China only act rarely as main compensators. China's share increases significantly until 2011, while Japan does not appear any more. Moreover, China's regional sectors are primarily main compensators in sectors of natural products (AGRI, FISH, etc.), but also extend its influence in sectors such as METL and MACH in 2011.



# 7 Network properties as main compensator predictors

## 7.1 Node property analysis

### Choice of properties

In order to better understand the main compensator behaviour of the regional sectors, ten node properties are defined to interpret the optimisation results:

$\deg_{ir}^{Z,in}$	Input degree of intermediate flows, Eq. (2.29)
$\deg_{ir}^{Z,out}$	Output degree of intermediate flows, Eq. (2.30)
$\deg_{ir}^Y$	Degree of final demand flows, Eq. (2.30)
$X_{ir}$	Total output, Eq. (2.27)
$D_{ir}$	Total input, Eq. (2.28)
$\alpha_{ir}$	Efficiency $\alpha_{ir} = 1 - D_{ir}/X_{ir}$
$N_{ir,i^*r^*}^{Z,in}$	Number of shared neighbours of intermediate input flows with the affected regional sector $i^*r^*$ , Eq. (2.37)
$N_{ir,i^*r^*}^{Z,out}$	Number of shared neighbours of intermediate output flows with the affected regional sector $i^*r^*$ , Eq. (2.38)
$N_{ir,i^*r^*}^{Y,out}$	Number of shared neighbours of final demand output flows with the affected regional sector $i^*r^*$ , Eq. (2.38)
$N_{ir,i^*r^*}^{out}$	Number of shared neighbours of output flows with the affected regional sector $i^*r^*$ , Eq. (2.38)

These properties represent topological features of the network as well as pure node properties. Especially the regional sector's total output property  $X$  is assumed to be of relevance, because the model's target function is minimising the relative output deviations. Moreover, the shared neighbour feature takes a specific simulation parameter  $i^*r^*$  (the affected regional sector) into account, which is of importance since the property signifies the connection to other regional sectors that are directly affected by the production loss of  $i^*r^*$ . These ten properties shall be used to bet-

## 7 Network properties as main compensator predictors

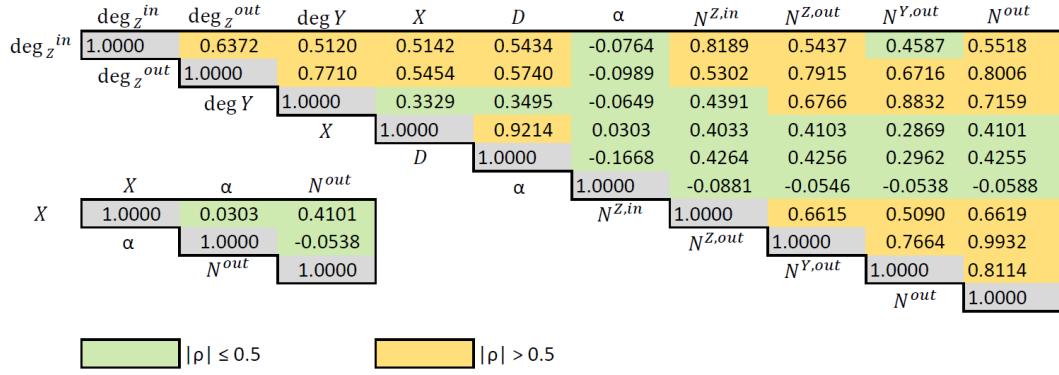


Figure 7.1: Averaged correlation coefficient matrix  $\rho$  and reduced matrix for selected node properties. The property  $\alpha$  is highly uncorrelated ( $|\rho| \leq 0.2$ ) to the other properties. Furthermore, the property  $N^{out}$  has low correlations ( $|\rho| \leq 0.5$ ) to  $X$  and  $D$ . Since these have a correlation coefficient of  $\rho_{X,D} = 0.9214$ , one of them can be omitted ( $D$ ). The remaining properties are:  $X$ ,  $N^{out}$  and  $\alpha$ .

ter understand the regional sectors' response to the exogenous forcing, interpreting them in the context of their production change  $p$ .

### Correlation analysis of properties

The regional sectors' main compensator behaviour shall be described through a function of independent properties. In order to ensure their independence, the correlation between the different properties is tested. Therefore, a correlation analysis is applied and a set of three least correlated properties is selected for further analysis. It is assumed that two properties are uncorrelated if the absolute value of their correlation coefficient is  $|\rho| < 0.5$ . The correlation matrix is obtained by averaging the correlation matrix of each feasible simulation run (with different  $i^*r^*$ ) of the year 2011. The results are shown in Fig. 7.1 along with the reduced correlation matrix of the selected weakly correlated properties  $X_{ir}$ ,  $N_{ir,i^*r^*}^{out}$  and  $\alpha_{ir}$ . The shared output neighbour property  $N_{ir,i^*r^*}^{out}$  was selected in favour of  $N_{ir,i^*r^*}^{Z,out}$  and  $N_{ir,i^*r^*}^{Y,out}$ , because it represents a more general node feature than the others do. It is denoted by  $N$  from now on.

### Regression vs. probability analysis

A potential method to connect the resulting production changes  $p$  with the nodes' properties would be a linear regression analysis. The optimisation results generally only have one or two regional sectors that change their production significantly, while the others stay roughly constant. This behaviour makes a linear regression

less useful, since there is no linear relation between change in production and node property ranking values. Instead, a binary relation formulation is favoured where a certain node with a given set of properties either changes its production significantly or not. Furthermore, it is possible that a node with a specific property combination (e.g. a relatively low value in  $X$  but a high value in  $\alpha$ ) shows a particular response behaviour that can not be related solely to the isolated ranking values of the properties, but rather to the relation of the properties to each other. This leads to conditional behaviour analysis, which is not appropriately expressible through a linear regression approach.

Therefore, a probability analysis is chosen instead of the regression analysis. The probability analysis aims at calculating the probability that an arbitrary regional sector with given properties does change its production significantly, taking the boolean behaviour into account. The concept of conditioned probabilities can be applied using *Bayes' theorem* (Papoulis, 1984):

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}, \quad (7.1)$$

where  $A$  and  $B$  are two events and  $P(A)$ ,  $P(B)$  the probabilities that these events occur, while  $P(A|B)$  is the probability that event  $A$  occurs under the condition that event  $B$  already did occur (the other way round for  $P(B|A)$ ).

### Ranking of node properties

In most simulation runs the main compensator behaviour is governed by the relative rank of a regional sector with respect to the other regional sectors, while the specific value of the node property is not. For instance, the regional sector with the highest value of  $X$  frequently acts as main compensator (as shown later in this chapter). For this reason, a node is ranked compared to the other nodes, where 1 stands for rank 1, meaning the regional sector is the strongest in this property.

The analysis of the three properties of an arbitrarily selected regional sector leads to an assessment of the event '*regional sector is strongest in this property*'. The property tuple

$$B = (B_X, B_N, B_\alpha) \quad (7.2)$$

contains the events  $B_X, B_N, B_\alpha$ , which are the ranking events of properties  $X, N$

and  $\alpha$ , that can take on the boolean values 'true' (T) or 'false' (F):

$$B_k = \begin{cases} F & \text{'regional sector is strongest in property } k \text{'} = \text{false}, \\ T & \text{'regional sector is strongest in property } k \text{'} = \text{true}. \end{cases} \quad (7.3)$$

The event tuple  $B$  contains information about the 'strongest in this property' events of all three properties.

## 7.2 Main compensator probability calculation

With the usage of the property event tuple  $B$  of an arbitrary regional sector within an arbitrary simulation run,  $P(A|B)$  is the probability that the regional sector will be the main compensator with the strongest change of  $p$ . Consider an arbitrary regional sector within any given simulation run, then the event  $A_j$  denotes if it is the main compensator, with  $j \in \{0, 1\}$  and

$$j = \begin{cases} 0 & \text{is not the main compensator,} \\ 1 & \text{is the main compensator} \end{cases} . \quad (7.4)$$

Using these definitions, we seek to analyse the probability  $P_i(A_1|B)$  that a property tuple  $B$  leads to a positive main compensator event  $A_1$  in a simulation with disrupted sector  $i = i^*$ . According to *Bayes' theorem* this probability is given by

$$P_i(A_1|B) = \frac{P_i(B|A_1) P_i(A_1)}{P_i(B)} . \quad (7.5)$$

**Term  $P_i(B|A_1)$**  describes the probability that a given main compensator (event  $A_1$ ) has a properties  $B$  in the case of a disruption in sector  $i$ . Since  $B$  is dependent on the simulation specifications, more precisely on the disrupted regional sector  $i^*r^*$  (property  $N$ ), each feasible simulation run  $\theta_t^i$  with a disruption in sector  $i = i^*$  is considered. Some simulation runs might be infeasible, hence only  $f_i \leq n_r$  feasible simulations of all  $n_r$  possible simulations are taken into account ( $t \in \{1, \dots, f_i\}$ ):

$$P_i(B|A_1) = \frac{1}{f_i} \sum_{t=1}^{f_i} \delta_{B, \hat{B}(\theta_t^i)} , \quad (7.6)$$

## 7.2 Main compensator probability calculation

with the delta function

$$\delta_{B,\hat{B}(\theta_t^i)} = \begin{cases} 1 & \text{if } B \text{ equals } \hat{B}(\theta_t^i), \\ 0 & \text{otherwise} \end{cases}, \quad (7.7)$$

where  $\hat{B}(\theta_t^i)$  denotes the properties of the main compensator in simulation  $\theta_t^i$ .

**Term  $P_i(A_1)$**  gives the probability that an arbitrary regional sector of sector  $i$  is a compensator in an arbitrary simulation run:

$$P_i(A_1) = \frac{1}{n_r - 1}. \quad (7.8)$$

Each regional sector of all  $n_r - 1$  available ones within a sector  $i$  could possibly be the main compensator (with the exception of  $i^*r^*$ ) with equally distributed probabilities for each one as long as no further information about node properties are given. Since each sector contains  $n_r$  regional sectors, the distinction between different sectors is not necessary:

$$P(A_1) = P_i(A_1) = \frac{1}{n_r - 1}, \forall i \in 1, \dots, n_s. \quad (7.9)$$

**Term  $P_i(B)$**  is the probability that an arbitrary regional sector of sector  $i$  has a property combination  $B$  within an arbitrarily selected simulation:

$$P_i(B) = \frac{1}{(n_r - 1) f_i} \sum_{t=1}^{f_i} n_i^B(\theta_t^i). \quad (7.10)$$

Here,  $n_i^B(\theta_t^i)$  denotes the amount of regional sectors producing good  $i$  and that have a property combination  $B$  for a given simulation run  $\theta_t^i$ .

**Reformulation of  $P_i(A_1|B)$**  leads to

$$P_i(A_1|B) = \frac{P_i(B|A_1) P(A_1)}{P_i(B)} \quad (7.11)$$

$$= \frac{(n_r - 1) f_i}{(n_r - 1) f_i} \frac{\sum_{t=1}^{f_i} \delta_{B,\hat{B}(\theta_t^i)}}{\sum_{t=1}^{f_i} n_i^B(\theta_t^i)} \quad (7.12)$$

$$= \frac{\sum_{t=1}^{f_i} \delta_{B,\hat{B}(\theta_t^i)}}{\sum_{t=1}^{f_i} n_i^B(\theta_t^i)} = \frac{n_{i,A_1}^B}{n_{i,tot}^B}, \quad (7.13)$$

which can be identified as the relation between the number of first compensators with properties  $B$  and the total number of regional sectors with properties  $B$ , both obtained from all possible simulations  $\{\theta^i\}$  with disruption in sector  $i$ .

### 7.3 Main compensator predictors

For a whole set of optimisation results of a given parameter set, we can determine the number of first compensators  $n_{i,A_1}^B$  with properties  $B$ . The total amount of regional sectors  $n_{i,tot}^B$  with property tuple  $B$  is obtained by analysing the aggregated EORA26 dataset. Using Eq. (7.13), the probabilities  $P_i(A_1|B)$  for each economic sector  $i$  are determined, which express how likely a regional sector with property tuple  $B$  acts as main compensator. The results are presented in a decision tree structure, where each of the possible property event results corresponds to one of two branches (left: 'regional sector is strongest in this property'=false, right: 'regional sector is strongest in this property'=true). While the leafs of a probability tree always add up to one, this is not the case for a decision tree. Instead, the complement rule  $P(A_1|B) + P(A_0|B) = 1$  holds. An exemplary decision tree is shown in Fig. 7.2.

We assess the ability of the 'strongest in property' events to act as a good main compensator predictors by calculating the measure  $\mu \in [0, 4]$  for each sector  $i$ :

$$\mu_i(X) = \sum_{\gamma=(T,F)} \sum_{\nu=(T,F)} P_i(A_1|(B_X = T, \gamma, \nu)), \quad (7.14)$$

$$\mu_i(N) = \sum_{\gamma=(T,F)} \sum_{\nu=(T,F)} P_i(A_1|(\gamma, B_N = T, \nu)), \quad (7.15)$$

$$\mu_i(\alpha) = \sum_{\gamma=(T,F)} \sum_{\nu=(T,F)} P_i(A_1|(\gamma, \nu, B_\alpha = T)). \quad (7.16)$$

This measure sums up all the probability values with a 'true' argument property event state. It must not be interpreted as a probability, but rather as a significance indicator.

An overview about all decision trees of year 2011 and the corresponding measures  $\mu$  for each sector is given in Fig. 7.4. Analogous calculations are applied to results for the years 2006, 2001 and 1996 (see Fig. A.10 - A.12).

The three properties have different impacts on the main compensator probability, dependent on the sector. In most of the cases, the ranking event of the total production output  $X$  determines if a regional sector acts as main compensator, but in some cases the shared output neighbours  $N$  and the efficiency  $\alpha$  also play an

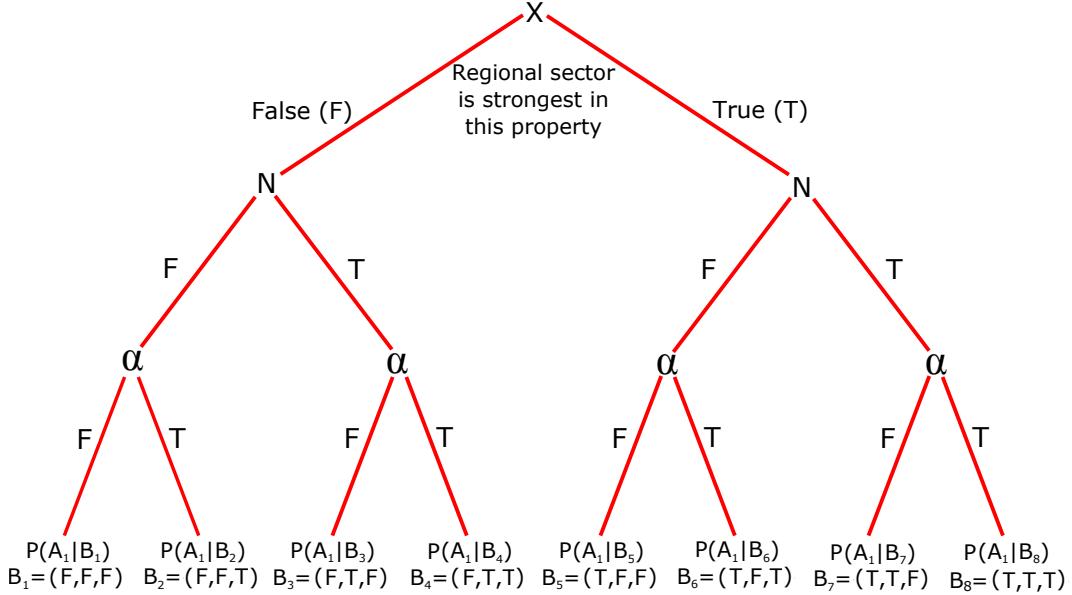


Figure 7.2: Exemplary decision tree for the property events  $B_\lambda$  and corresponding main compensator probabilities  $P(A_1|B_\lambda)$ ,  $\lambda \in \{1, \dots, 8\}$ . To determine  $P(A_1|B_\lambda)$  for a given regional sector with a known event set  $B_\lambda = (\{T, F\}, \{T, F\}, \{T, F\})$  that specifies the results of the property ranking events, one starts at the origin of the tree and determines whether the regional sector is strongest in property  $X$  and follows the respective branch. The procedure is repeated for the two remaining property events  $N$  and  $\alpha$ . The leafs of the tree show the probability values of  $P(A_1|B_\lambda)$ . While the leafs' probabilities of this decision tree do not necessarily add up to one, the complement rule  $P(A_1|B_\lambda) + P(A_0|B_\lambda) = 1$  holds.

important role (e.g. in the OILC and in the MACH sector). Furthermore, it is important to consider the specific values of the measure. The shared output neighbour predictor  $N$  for the TRAN sector has a measure of  $\mu(N) = 0.7$ , which is higher than  $\mu(X) = 0.4$  and  $\mu(\alpha) = 0$ , but the value is relatively low, because  $0 \leq \mu \leq 4$ . In this case, none of the predictors is significant. Contrarily, considering the FOOD sector, all predictors exhibit quite high measure values ( $\mu > 2$ ). Predicting the main compensator in this sector works well with all three 'strongest in property' events.

One limitation of the measure definition is the inability to assess 'false' event results and their impact on the probabilities. Taking a look at the TREQ sector, the event tuple  $B = (T, T, T)$  leads to a non-existent probability ('-'), which occurs due to a lack of data and can be interpreted as zero, while the event tuple  $B = (T, T, F)$  results in  $P(A_1|B) = 1$  (the asterisk can be omitted, since it indicates a rounded value). In this case, the property  $\alpha$  works as a good predictor, but in the sense that

## 7 Network properties as main compensator predictors

a ‘false’ event state leads to a high probability, what is not taken into account in the measure  $\mu$ .

Finally, we determine the property with the highest measure  $\mu$  for each sector and consider it to be the main predictor, assuming it predicts the main compensator best in the context of the optimisation results. We combine these main predictors with the main compensator assignments of chapter 6.3, the result is shown in Fig. 7.3. For the year 2011, China nearly exclusively acts as main compensator for sectors with the main predictor  $X$  (e.g. AGRI, FISH, MINQ, etc.), while about one third of the sectors with the USA as main compensator can be predicted well with the properties  $N$  and  $\alpha$  (e.g. OILC, CONS, TRAN, etc.). Taking a look at the three other years, one observes that the total output property  $X$  is nearly exclusively the main predictor.

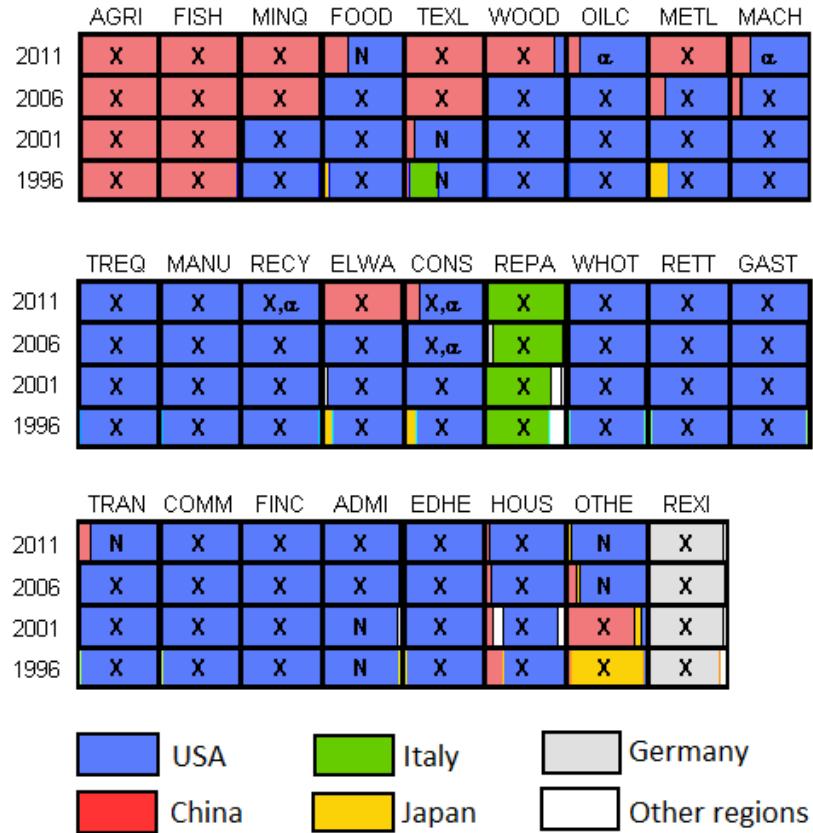


Figure 7.3: Share of regions in main compensators (colours) and main predictors ( $X, N, \alpha$ ) for each sector of the years 2011, 2006, 2001 and 1996. Results taken from simulation runs with parameters  $\epsilon = 0.5$  and using the LPG. In the case of identical measure values of two properties, both are considered to be main predictors.

### 7.3 Main compensator predictors

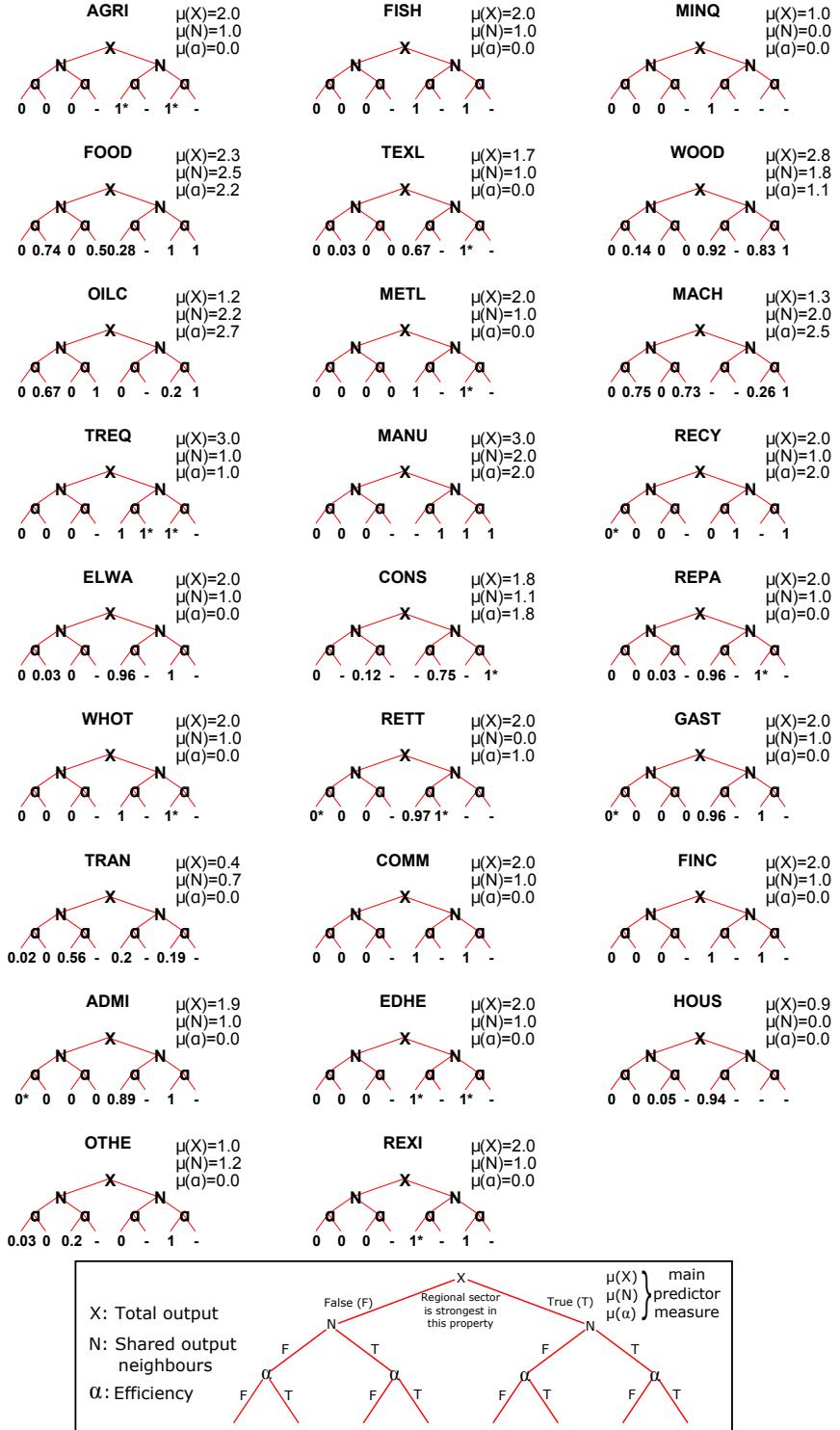


Figure 7.4: Decision tree of 2011 with  $\epsilon = 0.5$  for each sector with main compensator probabilities  $P(A_1|B)$  and corresponding measures  $\mu$ . The case of a non-existent probability (unavailability of necessary data) is represented by a hyphen '-' and can be interpreted as a probability of zero. An asterisk '\*' indicates that the probability was rounded to zero or one respectively.



# 8 Conclusions

## 8.1 Summary and Discussion

This thesis provides a framework to determine an optimal response of the global trade network to local production failures under specific constraint formulations, including analysis tools to investigate network properties and their impact on the optimisation results.

Due to the modular implementation of the software, it is easy to modify the optimisation problem's constraints and target function, allowing for an application of different adaptation mechanisms. In this thesis, an optimisation problem formulation with high adaptability (LPG) and an alternative one that reduces adaptation to flows of specific goods (LPS) were introduced. These two linear problems differ in their optimal responses, especially the target function takes higher values for the LPS due to less degrees of freedom for adaptation.

The value of the target function  $F(\mathbf{p})$  can be interpreted as the adaptation effort the world's trade network has to expend in order to compensate for the reduction of production in the affected regional sector  $i^*r^*$ . The adaptation effort for different failure locations  $F(i^*r^*)$ ,  $\forall i^*r^*$  shows high values for failures in China and the USA. Deeper insights about the structure of the results would be of interest, e.g. in the form of a case study in which the whole network is investigated thoroughly for a given affected regional sector  $i^*r^*$ .

For a large score of affected regional sectors, an infeasibility limit  $\tilde{\epsilon}_{i^*r^*}$  exists. In such a case, a feasible optimisation problem is given only for a failure strength  $\epsilon < \tilde{\epsilon}_{i^*r^*}$ . Above a critical value  $\epsilon > \epsilon_{c,i^*r^*}$ , that is specific to the failing regional sector, the adaptation effort  $F(\epsilon)$  shows a nonlinear behaviour. In this domain, the respective production ratios  $p_{ir}(\epsilon)$  seem to traverse several intervals in which they change linearly with different slopes for each interval.

Further investigation is necessary that clarifies if the existence of the different domains can be explained by network properties that could be subject of further research, or must rather be seen as a result of the simplex algorithm's numerical

## 8 Conclusions

errors. One hint for the first assumption is that most of the regional sectors with low values of  $\tilde{\epsilon}_{i^*r^*}$  have a high share in final demand flows. A disruption stronger than  $\tilde{\epsilon}_{i^*r^*}$  in a regional sector that is the main supplier of final demand of a specific region with only few alternative suppliers could result in an inability of the network to fulfill the constant final demand constraint. In such a case, the problem becomes infeasible. A more flexible final demand constraint with an adaptable consumption could reduce the occurrences of low values of  $\tilde{\epsilon}_{i^*r^*}$ .

In the optimisation results, usually regional sectors from China and the USA act as main compensators with the strongest increase of the production ratio  $p$  in the affected sector  $i^*$ , while all other regional sectors  $i^*r$  in general do not change their production ratio much (less than 10% of the main compensator's increase of  $p$ ). Problematic could be the fact that both countries represent a huge number of companies and a high population, what makes them less comparative to other countries in the world and could favour the role of their regional sectors as main compensators within our model.

Finally, three node characteristics that could be used as predictors for the main compensator were determined and analysed for different years. The most important insight is, that the main compensator for a failure in sector  $i^*$  frequently has the highest value in the total output property  $X$ , while the amount of shared output neighbours  $N$  with the affected regional sector  $i^*r^*$  and the efficiency  $\alpha$  of the regional sector play an important role for specific sectors. More detailed results could be acquired by taking more network properties into account, such as betweenness centrality or node degrees of higher order (taking several neighbour layers into account), with the aim of predicting the compensation behaviour more precisely.

All simulation runs were conducted with the aggregated EORA26 data set. Due to the aggregation of the data, some information inevitably got lost, reducing the accuracy of the simulation results. The aggregation process was necessary to reduce the complexity of the linear problem and hence the computing time.

During the implementation of the code a lot of obstacles occurred, especially due to the large number of variables and constraints. The predefined solving algorithms of MATLAB (MATLAB, 2012) were only capable of producing feasible solutions for problems of at least three magnitudes smaller than needed and were therefore not applicable. This led to a self-made implementation of the simplex algorithm with consideration of all relevant numerical errors<sup>1</sup>. The optimisation results have a high accuracy with a numerical equation system error  $\|Ax - b\|$  of  $\mathcal{O}(10^{-3})$  \$/year.

---

<sup>1</sup>Important for solving optimisation problems of this size is the consideration of the condition number (Belsley et al., 2004).

Another obstacle was the need for a large RAM, since the constraint matrix derived from the original EORA26 data exceeds the 4GB RAM limit of standard up-to-date PCs. The usage of special large memory nodes of the high performance cluster computers at PIK solved the issue. This resource problem shows that even by using supercomputers, solving optimisation problems of this size reaches the limits of today's technology, what may cause problems as soon as higher resolution data sets (as *zeean* is about to provide) become the center of research activities.

## 8.2 Outlook

The framework provided with this thesis offers a huge potential for further detailed investigation and the inclusion of sophisticated model extensions.

A finer resolution of regions would diminish the large differences in node properties between the USA and China on the one side and the rest of the countries on the other side, what might result in production adaptations also in other regions apart from China and the USA. The data set that *zeean* is about to provide could help out.

Another attempt to soften the main compensator effect would be the introduction of a quadratic target function. Large deviations from the initial production would be penalised disproportionately high, so that China and the USA would refrain from huge changes in production and other countries might act as compensators, too.

The assignment of the main compensators to sectors is another interesting topic for further research. The time evolution shows how China acts as main compensator in an increasing number of sectors. Especially a study with a connection to the product space (Hidalgo et al., 2007) would be of interest in order to better understand China's economical evolution. A higher resolution time series analysis for more years could show the path of the main compensators within the product space more accurately.

Further model expansions could be the inclusion of a rewiring rule in which new links could be established within the network, providing more flexibility and a more realistic adaptation behaviour.

To take the actual cost of producing the diverse goods  $i$  into account, the components of the cost vector  $\mathbf{c}$  of the target function could take on the corresponding costs for each sector, transforming the target function into a cost function. In this context, transportation costs could also be implemented.



# Appendix

## A.1 Sectors and regions

1	AGRI	Agriculture	14	CONS	Construction
2	FISH	Fishing	15	REPA	Maintenance and Repair
3	MINQ	Mining and Quarrying	16	WHOT	Wholesale Trade
4	FOOD	Food & Beverages	17	RETT	Retail Trade
5	TEXL	Textiles and Wearing Apparel	18	GAST	Hotels and Restaurants
6	WOOD	Wood and Paper	19	TRAN	Transport
7	OILC	Petroleum, Chemical and Non-Metallic Mineral Products	20	COMM	Post and Telecommunications
8	METL	Metal Products	21	FINC	Finacial Intermediation and Business Activities
9	MACH	Electrical and Machinery	22	ADMI	Public Administration
10	TREQ	Transport Equipment	23	EDHE	Education, Health and Other Services
11	MANU	Other Manufacturing	24	HOUS	Private Households
12	RECY	Recycling	25	OTHE	Others
13	ELWA	Electricity, Gas and Water	26	REXI	Re-export & Re-import

Figure A.1: Sectors, indices and abbreviations for EORA26 data set.

1	AFG	Afghanistan	39	CHL	Chile	77	HUN	Hungary	115	MAR	Morocco	153	SVK	Slovakia
2	ALB	Albania	40	CHN	China	78	ISL	Iceland	116	MOZ	Mozambique	154	SVN	Slovenia
3	DZA	Algeria	41	COL	Colombia	79	IND	India	117	MMR	Myanmar	155	SOM	Somalia
4	AND	Andorra	42	COG	PR Congo	80	IDN	Indonesia	118	NAM	Namibia	156	ZAF	South Africa
5	AGO	Angola	43	CRI	Costa Rica	81	IRN	Iran	119	NPL	Nepal	157	ESP	Spain
6	ATO	Antigua and Barbuda	44	HRV	Croatia	82	IRQ	Iraq	120	NLD	Netherlands	158	LKA	Sri Lanka
7	ARG	Argentina	45	CUB	Cuba	83	IRL	Ireland	121	ANT	Netherlands Antilles	159	SUR	Surinam
8	ARM	Armenia	46	CYP	Cyprus	84	ISR	Israel	122	NCL	New Caledonia	160	SWZ	Swaziland
9	ABW	Aruba	47	CZE	Czech Republic	85	ITA	Italy	123	NZL	New Zealand	161	SWE	Sweden
10	AUS	Australia	48	CIV	Côte d'Ivoire	86	JAM	Jamaica	124	NIC	Nicaragua	162	CHE	Switzerland
11	AUT	Austria	49	PRK	North Corea	87	JPN	Japan	125	NER	Niger	163	SYR	Syria
12	AZE	Azerbaijan	50	COD	DR Congo	88	JOR	Jordan	126	NGA	Nigeria	164	TWN	Taiwan
13	BHS	Bahamas	51	DNK	Denmark	89	KAZ	Kazakhstan	127	NOR	Norway	165	TJK	Tajikistan
14	BHR	Bahrain	52	DJI	Djibouti	90	KEN	Kenya	128	PSE	Palestine	166	THA	Thailand
15	BGD	Bangladesh	53	DOM	Dominican Republic	91	KWT	Kuwait	129	OMN	Oman	167	MKD	Macedonia
16	BRB	Barbados	54	ECU	Ecuador	92	KGZ	Kyrgyzstan	130	PAK	Pakistan	168	TGO	Togo
17	BLR	Belarus	55	EGY	Egypt	93	LAO	Laos	131	PAN	Panama	169	TTO	Trinidad and Tobago
18	BEL	Belgium	56	SLV	El Salvador	94	LVA	Latvia	132	PNG	Papua New Guinea	170	TUN	Tunisia
19	BLZ	Belize	57	ERI	Eritrea	95	LBN	Lebanon	133	PRY	Paraguay	171	TUR	Turkey
20	BEN	Benin	58	EST	Estonia	96	LSO	Lesotho	134	PER	Peru	172	TKM	Turkmenistan
21	BMU	Bermuda	59	ETH	Ethiopia	97	LBR	Liberia	135	PHL	Philippines	173	UGA	Uganda
22	BTN	Bhutan	60	FJI	Fiji	98	LIB	Libya	136	POL	Poland	174	UKR	Ukraine
23	BOL	Bolivia	61	FIN	Finland	99	LIE	Liechtenstein	137	PRT	Portugal	175	ARE	United Arab Emirates
24	BIH	Bosnia and Herzegovina	62	FRA	France	100	LTU	Lithuania	138	QAT	Qatar	176	GBR	United Kingdom
25	BWA	Botswana	63	PYF	French Polynesia	101	LUX	Luxembourg	139	KOR	South Korea	177	TZA	Tanzania
26	BRA	Brazil	64	GAB	Gabon	102	MAC	Macao	140	MDA	Moldova	178	USA	United States of America
27	VGB	British Virgin Islands	65	GMB	Gambia	103	MDG	Madagascar	141	ROU	Romania	179	URY	Uruguay
28	BRN	Brunei	66	GEO	Georgia	104	MWI	Malawi	142	RUS	Russia	180	UZB	Uzbekistan
29	BGR	Bulgaria	67	DEU	Germany	105	MYS	Malaysia	143	RWA	Rwanda	181	VUT	Vanuatu
30	BFA	Burkina Faso	68	GHA	Ghana	106	MDV	Maldives	144	WSM	Samoa	182	VEN	Venezuela
31	BDI	Burundi	69	GRC	Greece	107	MLI	Mali	145	SMR	San Marino	183	VNM	Vietnam
32	KHM	Cambodia	70	GRL	Greenland	108	MLT	Malta	146	STP	Sao Tome and Principe	184	YEM	Yemen
33	CMR	Cameroon	71	GTM	Guatemala	109	MRT	Mauritania	147	SAU	Saudi Arabia	185	ZMB	Zambia
34	CAN	Canada	72	GIN	Guinea	110	MUS	Mauritius	148	SEN	Senegal	186	ZWE	Zimbabwe
35	CPV	Cap Verde	73	GUY	Guyana	111	MEX	Mexico	149	SRB	Serbia			
36	CYM	Cayman Islands	74	HTI	Haiti	112	MCO	Monaco	150	SYC	Seychelles			
37	CAF	Central African Republic	75	HND	Honduras	113	MNG	Mongolia	151	SLE	Sierra Leone			
38	TCD	Chad	76	HKG	Hong Kong	114	MNE	Montenegro	152	SGP	Singapore			

Figure A.2: Indices, official UN country codes and names of countries for EORA26 data set.

## Appendix

	single regions	29 rest of Europe	30 rest of Asia	31 rest of America	32 rest of Africa
1	Argentina	Albania	Afghanistan	Antigua and Barbuda	Algeria
2	Australia	Andorra	Armenia	Aruba	Angola
3	Brazil	Austria	Azerbaijan	Bahamas	Benin
4	Canada	Belarus	Bahrain	Barbados	Botswana
5	China	Belgium	Bangladesh	Belize	Burkina Faso
6	France	Bosnia and Herzegovina	Bhutan	Bermuda	Burundi
7	Germany	Bulgaria	Brunei Darussalam	Bolivia	Cameroon
8	India	Croatia	Cambodia	British Virgin Islands	Cape Verde
9	Indonesia	Cyprus	North Corea	Cayman Islands	Central African Republic
10	Italy	Czech Republic	Fiji	Chile	Chad
11	Japan	Denmark	French Polynesia	Colombia	Congo
12	Mexico	Estonia	Georgia	Costa Rica	Cote d'Ivoire
13	Netherlands	Finland	Hong Kong	Cuba	Congo
14	New Zealand	Greece	Iran	Dominican Republic	Djibouti
15	Nigeria	Greenland	Iraq	Ecuador	Egypt
16	Norway	Hungary	Israel	El Salvador	Eritrea
17	Poland	Iceland	Jordan	Guatemala	Ethiopia
18	Republic of Korea	Ireland	Kazakhstan	Guyana	Gabon
19	Russian Federation	Latvia	Kuwait	Haiti	Gambia
20	Saudi Arabia	Liechtenstein	Kyrgyzstan	Honduras	Ghana
21	South Africa	Lithuania	Laos	Jamaica	Guinea
22	Spain	Luxembourg	Lebanon	Netherlands Antilles	Kenya
23	Sweden	Malta	Macau	Nicaragua	Lesotho
24	Switzerland	Monaco	Malaysia	Panama	Liberia
25	Turkey	Montenegro	Maldives	Paraguay	Libyan Arab Jamahiriya
26	United Arab Emirates	Portugal	Mongolia	Peru	Madagascar
27	UK	Moldova	Myanmar	Suriname	Malawi
28	USA	Romania	Nepal	Trinidad and Tobago	Mali
		San Marino	New Caledonia	Uruguay	Mauritania
		Serbia	Palestine	Venezuela	Mauritius
		Slovakia	Oman		Morocco
		Slovenia	Pakistan		Mozambique
		Macedonia	Papua New Guinea		Namibia
		Ukraine	Philippines		Niger
			Qatar		Rawanda
			Samoa		Sao Tome and Principe
			Singapore		Senegal
			Sri Lanka		Seychelles
			Syrian Arab Republic		Sierra Leone
			Taiwan		Somalia
			Tajikistan		Swaziland
			Thailand		Togo
			Turkmenistan		Tunisia
			Uzbekistan		Uganda
			Vanuatu		Tanzania
			Vietnam		Zambia
			Yemen		Zimbabwe

Figure A.3: Indices and names of the regions in the aggregated EORA26 database. The 28 single countries remain distinct regions, while the other countries are aggregated to four proxy *rest of* regions, resulting in 32 regions, instead of the original 186.

## A.2 Network adaptation effort $F(i^*r^*)$

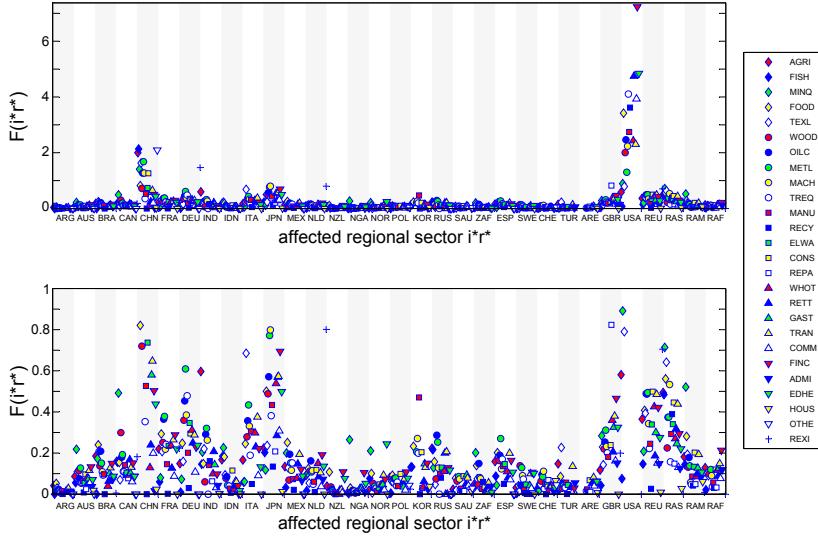


Figure A.4: Network adaptation effort  $F(i^*r^*)$  of all feasible simulation runs within the linear domain for the year 2006 and a forcing of  $\epsilon = 0.5$ , using the LPG. Compared to the values of 2011, China has a much smaller impact on global trade.

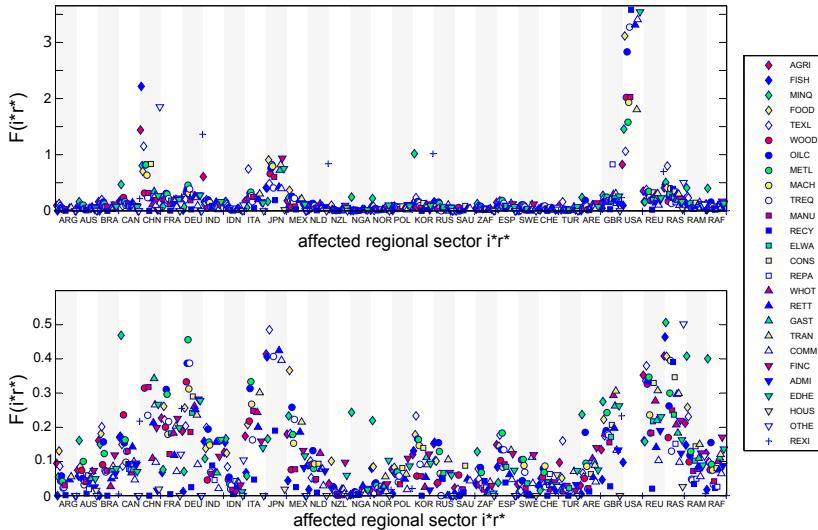


Figure A.5: Network adaptation effort  $F(i^*r^*)$  of all feasible simulation runs within the linear domain for the year 2001 and a forcing of  $\epsilon = 0.5$ , using the LPG. Japan exhibits relatively large values compared to later years.

## Appendix

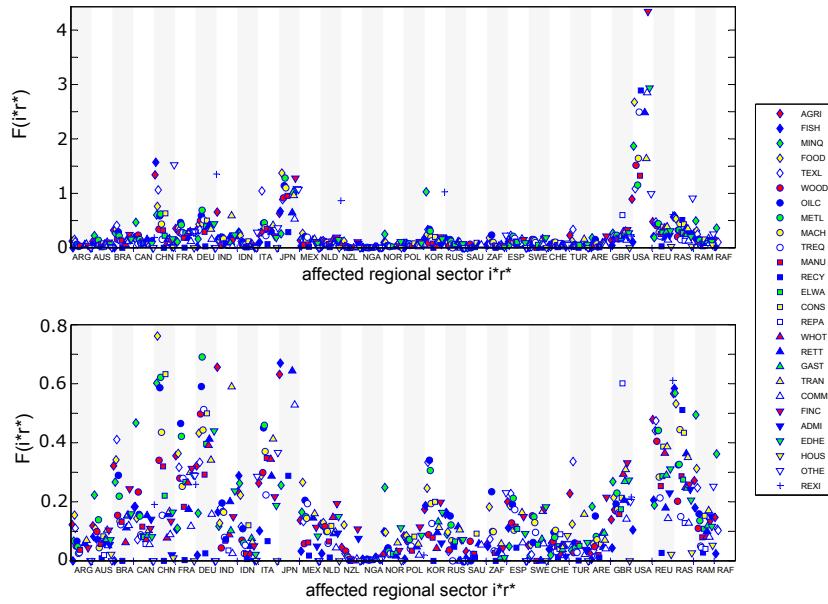


Figure A.6: Network adaptation effort  $F(i^*r^*)$  of all feasible simulation runs within the linear domain for the year 1996 and a forcing of  $\epsilon = 0.5$ , using the LPG. China and Japan both show high values, while also Germany, France and Korea exhibit high adaptation efforts.

### A.3 Main compensator overviews

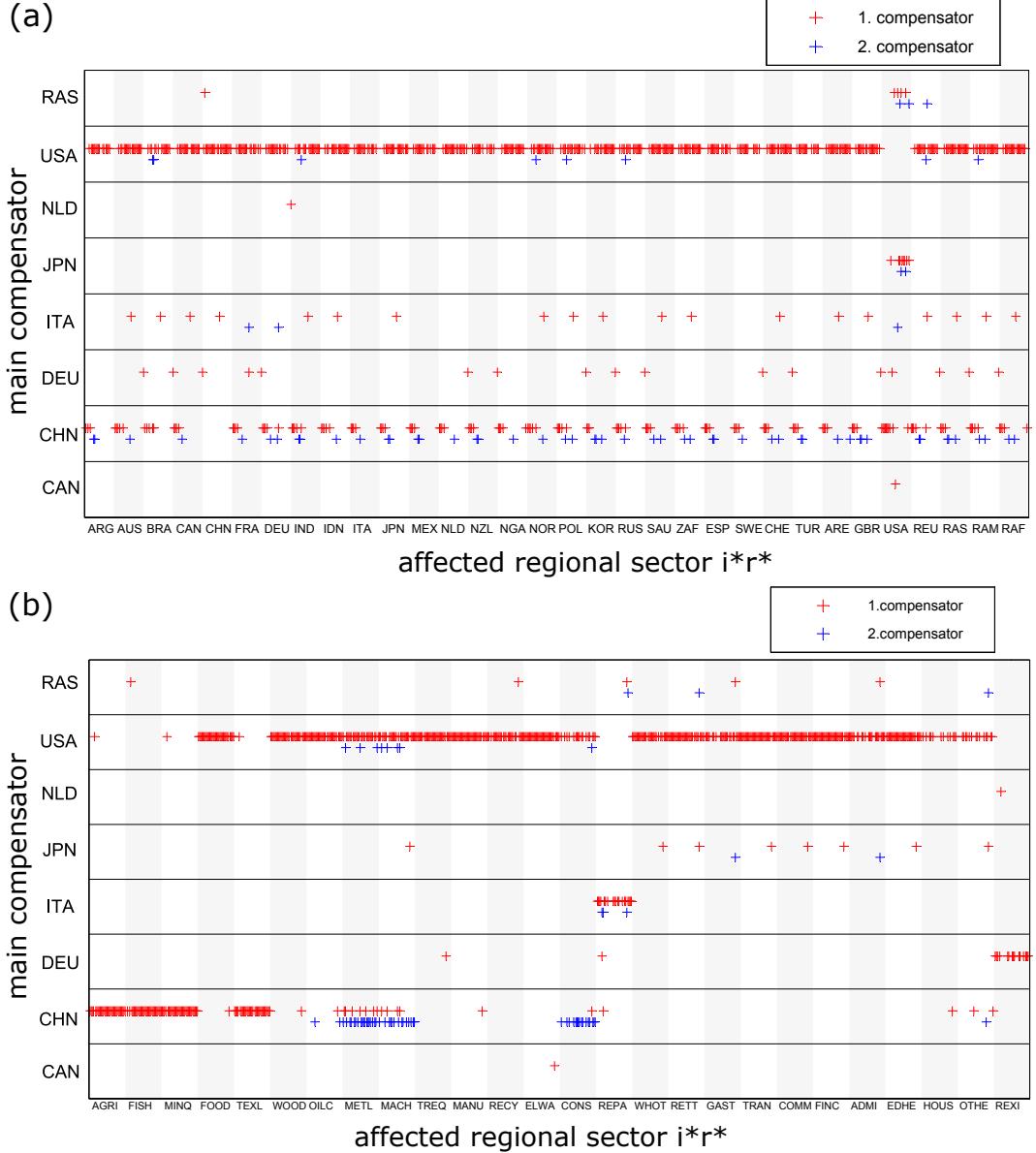


Figure A.7: (a) Overview of the main compensators sorted by regions for the year 2006,  $\epsilon = 0.5$ , using the LPG and the aggregated EORA26 data, where each feasible scenario with affected regional sector  $i^*r^*$  is displayed on the x axis. In most of the scenarios, China or the USA act as main compensators, while some other regions act as main compensators in specific single scenarios. (b) Same overview, sorted by sectors. It is apparent that China acts as main compensator for less sectors, compared to 2011.

## Appendix

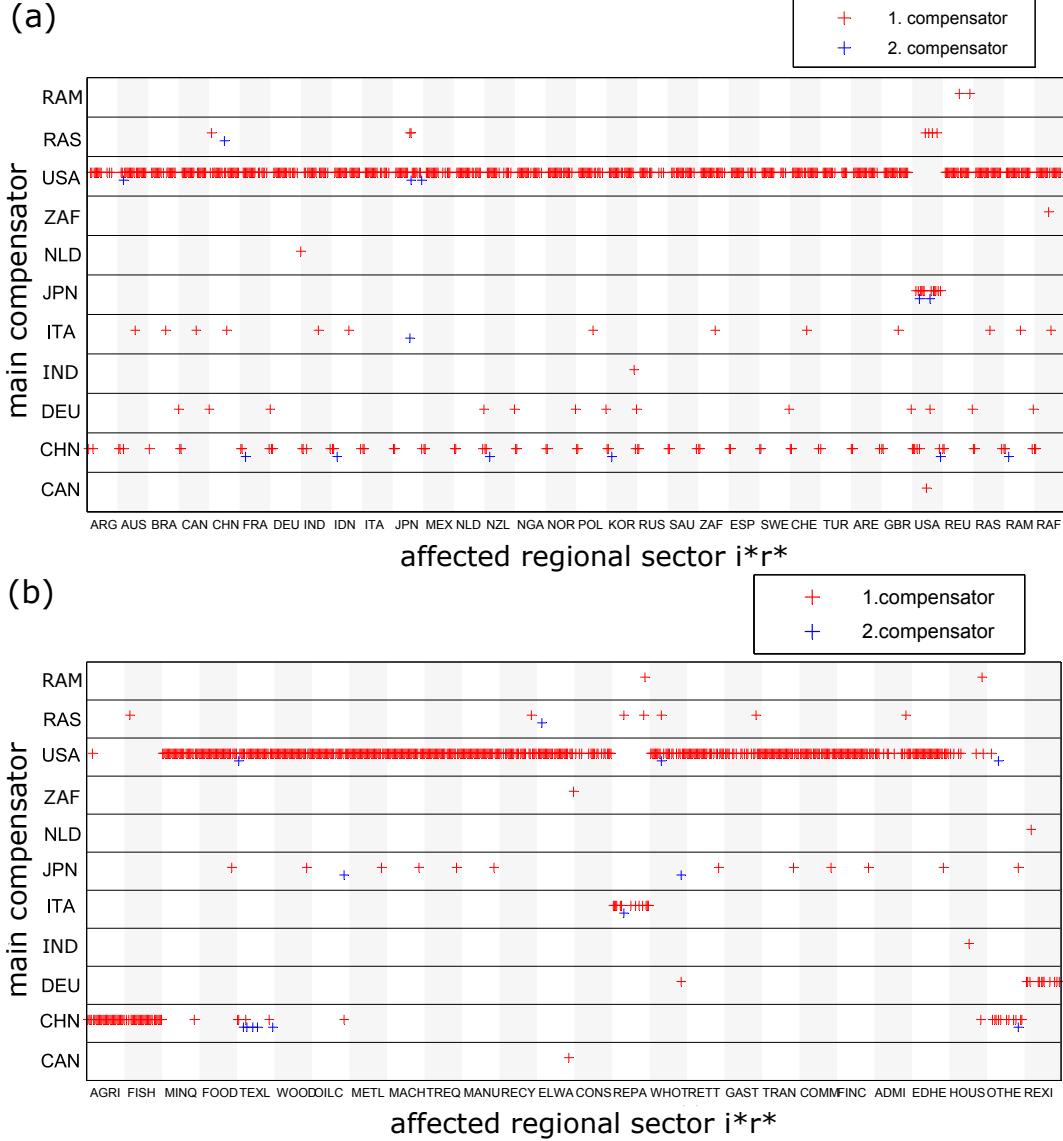


Figure A.8: (a) Overview of the main compensators sorted by regions for the year 2001,  $\epsilon = 0.5$ , using the LPG and the aggregated EORA26 data, where each feasible scenario with affected regional sector  $i^*r^*$  is displayed on the x axis. In most of the scenarios, the USA act as main compensators, while some other regions act as main compensators in specific single scenarios. (b) Same overview, sorted by sectors. China acts as main compensator for only three sectors, compared to 7 in 2011.

### A.3 Main compensator overviews

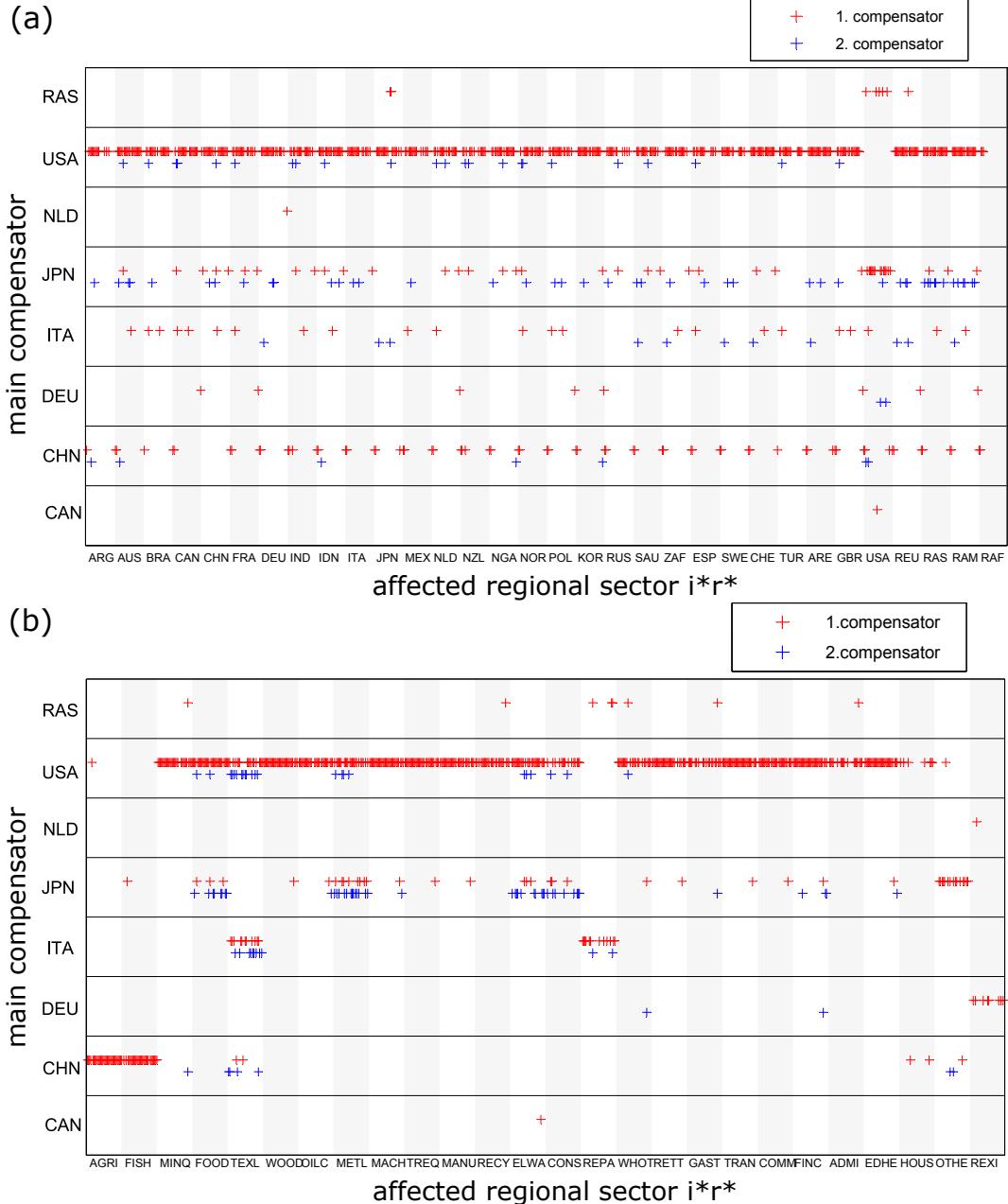


Figure A.9: (a) Overview of the main compensators sorted by regions for the year 1996,  $\epsilon = 0.5$ , using the LPG and the aggregated EORA26 data, where each feasible scenario with affected regional sector  $i^*r^*$  is displayed on the x axis. The USA act predominantly as main compensators in most of the sectors, while China is strong in the AGRI and FISH sectors. In contrast to later years, Japan is the main compensator quite frequently and Italy compensates partly in the TEXL sector. (b) Same overview, sorted by sectors.

## Appendix

### A.4 Main compensator probabilities

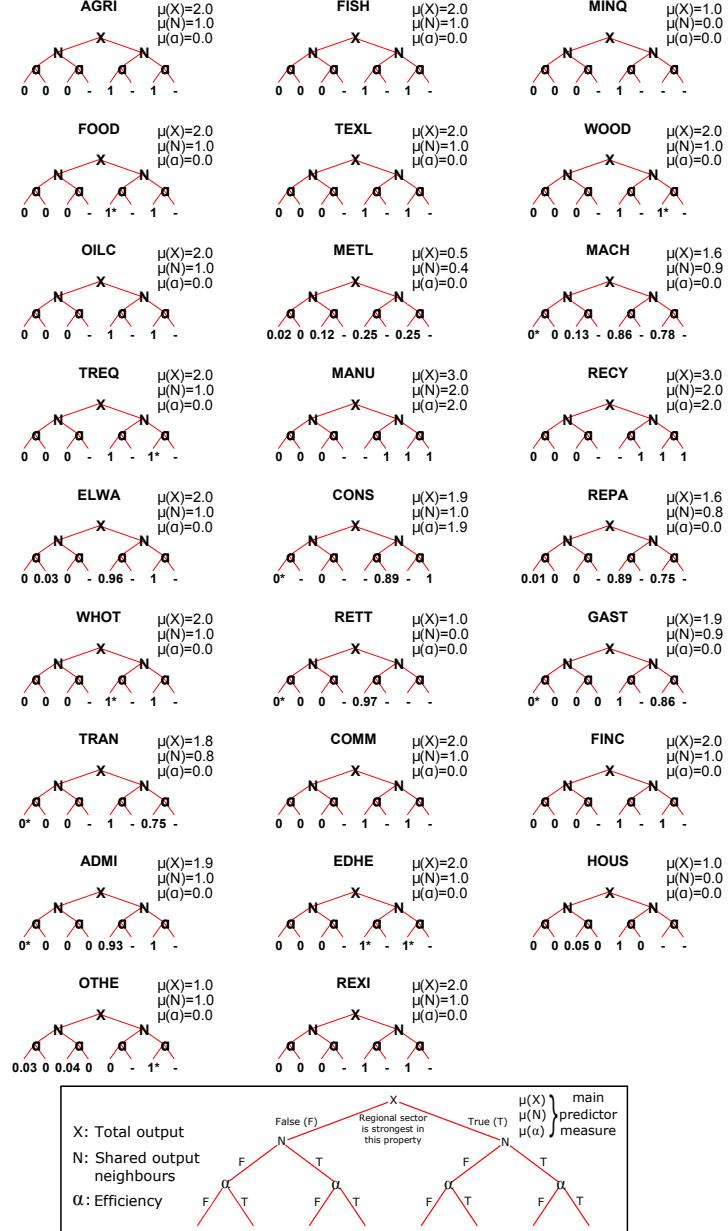


Figure A.10: Decision tree of 2006 with  $\epsilon = 0.5$  for each sector with main compensator probabilities  $p(A_1|B)$  and corresponding measures  $\mu$ . The case of a non-existent probability (unavailability of necessary data) is represented by a hyphen '-' and can be interpreted as a probability of zero. An asterisk '\*' indicates that the probability was rounded to zero or one respectively.

#### A.4 Main compensator probabilities

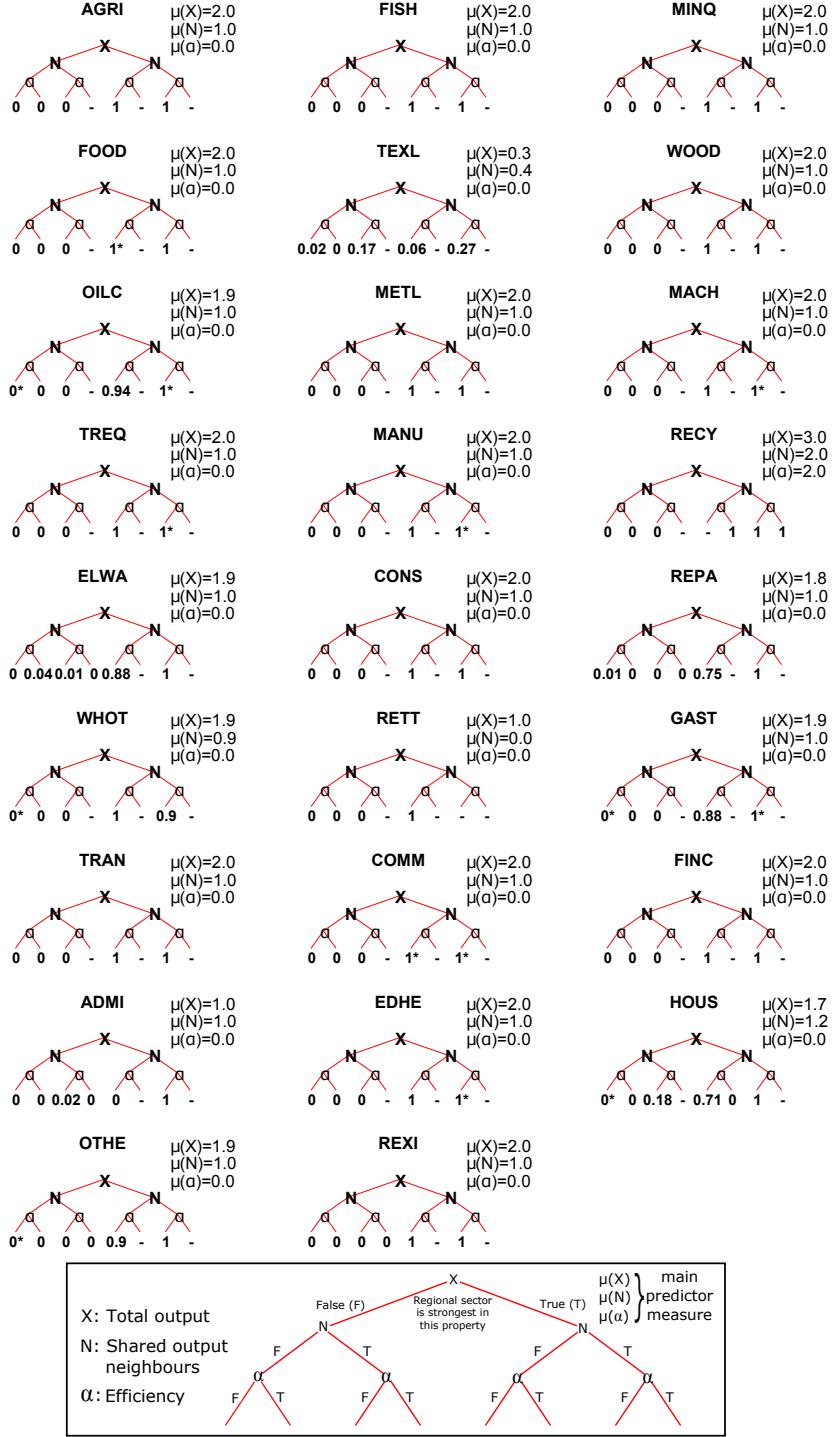


Figure A.11: Decision tree of 2001 with  $\epsilon = 0.5$  for each sector with main compensator probabilities  $p(A_1|B)$  and corresponding measures  $\mu$ . The case of a non-existent probability (unavailability of necessary data) is represented by a hyphen '-' and can be interpreted as a probability of zero. An asterisk '\*' indicates that the probability was rounded to zero or one respectively.

Appendix

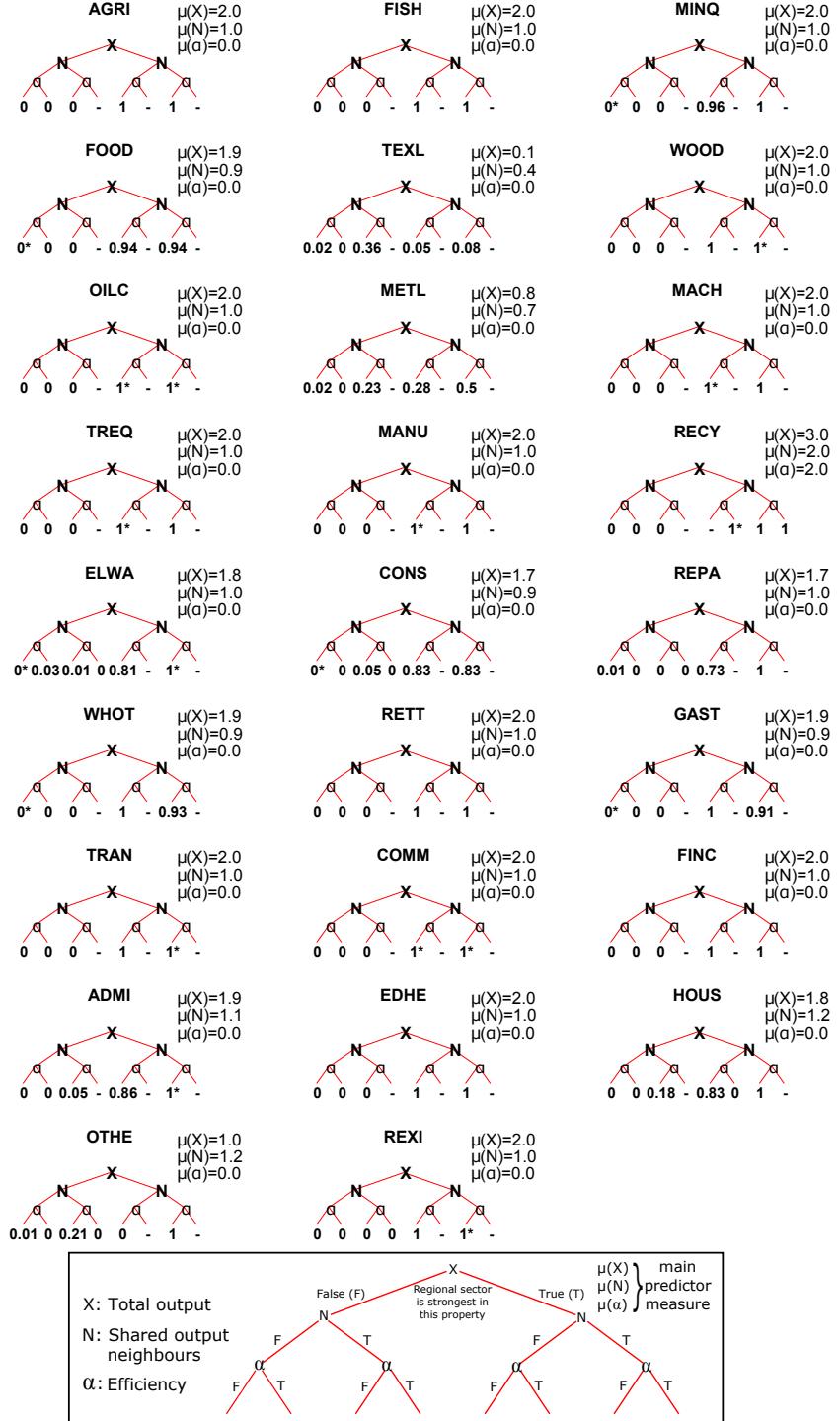


Figure A.12: Decision tree of 1996 with  $\epsilon = 0.5$  for each sector with main compensator probabilities  $p(A_1|B)$  and corresponding measures  $\mu$ . The case of a non-existent probability (unavailability of necessary data) is represented by a hyphen '-' and can be interpreted as a probability of zero. An asterisk '\*' indicates that the probability was rounded to zero or one respectively.

# Bibliography

- (Algarin et al., 2015) Algarin, J. V., Hawkins, T. R., Marriott, J., Matthews, H. S., and Khanna, V. (2015). *Disaggregating the Power Generation Sector for Input-Output Life Cycle Assessment*. Journal of Industrial Ecology, 19(4):666–675.
- (Allen, 1968) Allen, R. G. D. (1968). *Macro-economic Theory: A Mathematical Treatment*, page 35. Macmillan and St. Martin's Press, New York.
- (Beamon, 1998) Beamon, B. M. (1998). *Supply Chain Design and Analysis: Models and Methods*. International Journal of Production Economics, 55:281–294.
- (Belsley et al., 2004) Belsley, D. A., Kuh, E., and Welsch, R. E. (2004). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, chapter The Condition Number, pages 100–104. John Wiley & Sons, Hoboken, New Jersey.
- (Berrittella et al., 2006) Berrittella, M., Bigano, A., Roson, R., and Tol, R. S. J. (2006). *A general equilibrium analysis of climate change impacts on tourism*. Tourism Management, 27(5):913 – 924.
- (Bertsimas and Tsitsiklis, 1997) Bertsimas, D. and Tsitsiklis, J. N. (1997). *Introduction to Linear Optimization*, volume Book 6 of *Athena Scientific Series in Optimization and Neural Computation*. Athena Scientific, Nashua, third edition.
- (Bierkandt et al., 2014) Bierkandt, R., Wenz, L., Willner, S. N., and Levermann, A. (2014). *Acclimate - a model for economic damage propagation. Part I: basic formulation of damage transfer within a global supply network and damage conserving dynamics*. Environment Systems and Decisions, 34(4):507–524.
- (Crowther et al., 2007) Crowther, K. G., Haimes, Y. Y., and Taub, G. (2007). *Systemic Valuation of Strategic Preparedness Through Application of the Inoperability Input-Output Model with Lessons Learned from Hurricane Katrina*. Risk Analysis, 27(5):1345–1364.
- (Dantzig, 1987) Dantzig, G. B. (1987). *Origins of the simplex method*. Systems Optimization Laboratory, 87(5).

## Bibliography

- (Eboli et al., 2010) Eboli, F., Parrado, R., and Roson, R. (2010). *Climate-change feedback on economic growth: explorations with a dynamic general equilibrium model.* null, 15:515–533.
- (Escaith, 2009) Escaith, H. (2009). *Trade Collapse, Trade Relapse and Global Production Networks: Supply Chains in the Great Recession.* World Trade Organization (WTO).
- (Haimes et al., 2008) Haimes, Y., Santos, J., Crowther, K., Henry, M., Lian, C., and Yan, Z. (2008). *Critical Infrastructure Protection*, chapter Risk Analysis in Interdependent Infrastructures, pages 297–310. Springer US, Boston, MA.
- (Hallegatte, 2008) Hallegatte, S. (2008). *An Adaptive Regional Input-Output Model and its Application to the Assessment of the Economic Cost of Katrina.* Risk Analysis, 28(3):779–799.
- (Hertel, 1997) Hertel, T. (1997). *Global Trade Analysis: Modeling and Applications.* Cambridge University Press, Cambridge.
- (Hidalgo et al., 2007) Hidalgo, C. A., Klinger, B., Barabási, A.-L., and Hausmann, R. (2007). *The Product Space Conditions the Development of Nations.* Science, 317(5837):482–487.
- (IDE-JETRO, 2006) IDE-JETRO (2006). *How to make Asian Input-Output Tables.* Institute of Developing Economies / JETRO.
- (IPCC, 2012) IPCC (2012). *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (SREX).* [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, UK, and New York, NY, USA, 582 pp.
- (Lenzen et al., 2012) Lenzen, M., Kanemoto, K., Moran, D., and Geschke, A. (2012). *Mapping the Structure of the World Economy.* Environmental Science & Technology, 46(15):8374–8381.
- (Lenzen et al., 2013) Lenzen, M., Moran, D., Kanemoto, K., and Geschke, A. (2013). *Building Eora: A Global Multi-Region Input-Output Database At High Country and Sector Resolution.* Economic Systems Research, 25(1):20–49.

## Bibliography

- (Leontief, 1986) Leontief, W. (1986). *Input-Output Economics*. Oxford University Press, Oxford, second edition.
- (Levermann, 2014) Levermann, A. (2014). *Climate economics: Make supply chains climate-smart*. Nature, 506:27–29.
- (MacKenzie et al., 2012) MacKenzie, C. A., Santos, J. R., and Barker, K. (2012). *Measuring changes in international production from a disruption: Case study of the Japanese earthquake and tsunami*. International Journal of Production Economics, 138.
- (Maluck and Donner, 2015) Maluck, J. and Donner, R. V. (2015). *A Network of Networks Perspective on Global Trade*. PLoS ONE, 10(7):1–24.
- (MATLAB, 2012) MATLAB (2012). R2012b. The MathWorks Inc., Natick, Massachusetts.
- (Mechler and Bouwer, 2015) Mechler, R. and Bouwer, L. M. (2015). *Understanding trends and projections of disaster losses and climate change: is vulnerability the missing link?* Climate Change, 133(1):23–35.
- (Nagurney et al., 2002) Nagurney, A., Dong, J., and Zhang, D. (2002). *A supply chain network equilibrium model*. Transportation Research Part E: Logistics and Transportation Review, 38(5):281 – 303.
- (Narayanan et al., 2015) Narayanan, G., Badri, A. A., and McDougall, R. (2015). *Global Trade, Assistance, and Production: The GTAP 9 Data Base*. Center for Global Trade Analysis, Purdue University.
- (Newman, 2010) Newman, M. (2010). *Networks: An Introduction*. Oxford University Press, Oxford.
- (Nordhaus and Yang, 1996) Nordhaus, W. D. and Yang, Z. (1996). *A Regional Dynamic General-Equilibrium Model of Alternative Climate-Change Strategies*. The American Economic Review, 86(4):741–765.
- (Okuyama et al., 2004) Okuyama, Y., Hewings, G. J. D., and Sonis, M. (2004). *Modeling Spatial and Economic Impacts of Disasters*, chapter Measuring Economic Impacts of Disasters: Interregional Input-Output Analysis Using Sequential Interindustry Model, pages 77–101. Springer Berlin Heidelberg, Berlin, Heidelberg.

## Bibliography

- (Otto et al., 2016) Otto, C., Willner, S. N., Wenz, L., Frieler, K., and Levermann, A. (2016). *Acclimate - a dynamic model for economic loss-propagation*. submitted: Economic Modelling.
- (Papoulis, 1984) Papoulis, A. (1984). *Probability, Random Variables, and Stochastic Processes*, chapter Bayes' Theorem in Statistics, pages 38–39. McGraw-Hill Inc., New York.
- (Press et al., 2002) Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2002). *Numerical Recipes in C++: The Art of Scientific Computing*. Cambridge University Press, Cambridge, second edition.
- (Rahmstorf and Coumou, 2011) Rahmstorf, S. and Coumou, D. (2011). *Increase of extreme events in a warming world*. Proceedings of the National Academy of Sciences, 108(44):17905–17909.
- (Samuelson and Nordhaus, 2004) Samuelson, P. A. and Nordhaus, W. D. (2004). *Economics*. McGraw-Hill, Columbus, OH, 18 edition.
- (Schulte in den Bäumen et al., 2015) Schulte in den Bäumen, H., Többen, J., and Lenzen, M. (2015). *Labour forced impacts and production losses due to the 2013 flood in Germany*. Journal of Hydrology, 527:142–150.
- (Timmer et al., 2015) Timmer, M., Dietzenbacher, E., Los, B., Stehrer, R., and de Vries, G. J. (2015). *An Illustrated User Guide to the World Input-Output Database: the Case of Global Automotive Production*. Review of International Economics, 23:575–605.
- (Troy et al., 2015) Troy, T. J., Kipgen, C., and Pal, I. (2015). *The impact of climate extremes and irrigation on US crop yields*. Environ. Res. Lett., 10:054013.
- (Tukker et al., 2009) Tukker, A., Poliakov, E., Heijungs, R., Hawkins, T., Neuwahl, F., Rueda-Cantuche, J., Giljum, S., Moll, S., Oosterhaven, J., and Bouwmeester, M. (2009). *Towards a global multi-regional environmentally extended input-output database*. Ecological Economics, 68(7):1928–1937.
- (United Nations, 2008) United Nations (2008). *International Standard Industrial Classification of All Economic Activities (ISIC)*. United Nations Publication, New York.

## Bibliography

- (van der Veen, 2004) van der Veen, A. (2004). *Disasters and economic damage: macro, meso and micro approaches.* Disaster Prevention and Management, 13:274–279.
- (Wenz et al., 2014) Wenz, L., Willner, S. N., Bierkandt, R., and Levermann, A. (2014). *Acclimate - a model for economic damage propagation. Part II: a dynamic formulation of the backward effects of disaster-induced production failures in the global supply network.* Environment Systems and Decisions, 34(4):525–539.
- (Wiedmann et al., 2011) Wiedmann, T., Wilting, H. C., Lenzen, M., Lutter, S., and Palm, V. (2011). *Quo Vadis MRIO? Methodological, data and institutional requirements for multi-region input-output analysis.* Ecological Economics, 70(11):1937–1945.



# Acknowledgements

I would like to thank the people who supported me in all development stages of this thesis.

First, I thank Anders Levermann and Leonie Wenz, who gave me the opportunity to do research in an interdisciplinary field combining economics, network theory and numerical modelling at the *Potsdam Institute for Climate Impact Research*. Both supported me with a lot of professional and educational advice and several fruitful discussions. Especially the regular meetings with Leonie with intense discussion and a lot of feedback were particularly important to me. Thanks also go to Nicole Glanemann for being my second referee and giving me insights into research fields related to my thesis.

Sven Willner and Christian Otto supported me at important pivot points, contributing fresh opinions and new perspectives. Without Stefan Schinkel I would not have been able to use all these smart Linux tools that make a programmer's life so much easier - thanks for that. In general, I want to thank the *Global Adaptation Strategies* group for the pleasant atmosphere and fun lunch times together. Furthermore, the assistance of the group's coordinators and the fast replies of the PIK's IT staff to the scores of inquiries I heaped on them were of great value to me.

Janjenka Szillat, Jan Wohland, Timo Reckling and Leonie Wenz were so kind to proofread this thesis. Lots of helpful comments found their way into the final version. Thanks also goes to the Richard-Winter-Stiftung for supporting my studies financially and giving me the freedom to concentrate entirely on my research.

Finally, I want to thank my parents and my friends for supporting me in the paths I choose to take.

**Erklärung**

Hiermit erkläre ich, dass ich diese Abschlussarbeit selbstständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht habe.

Darüber hinaus erkläre ich, dass diese Abschlussarbeit nicht, auch nicht auszugsweise, im Rahmen einer nichtbestandenen Prüfung an dieser oder einer anderen Hochschule eingereicht wurde.

Potsdam, den 21. April 2016

(Sebastian Klipp)