Modelling global human systems using a large-scale economic model, with some novel approaches to the analysis and visualisation of its outputs

Robert Graham Levy

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
of
University College London.

Centre for Advanced Spatial Analysis
University College London

2018

I, Robert Graham Levy, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

The work of this thesis was conducted with the financial support of the Engineering and Physical Sciences Research Council (EPSRC) under the grant ENFOLDing[1].

The structure of the model described in chapter 3 was developed in discussion with Thomas P. Oléron Evans. I declare that the writing and other research of this chapter, except the ideas for the structure of the model which were developed together, was solely my own work.

The work of chapter 7 is based on ideas discussed with Robert J. Downes, in particular the mathematical appendix of section 7.A on p. 203 in which most of the mathematics is his. I declare that the research and all other content except this section was solely my own work. None of the work of this chapter has appeared elsewhere.

[1] Explaining, Modelling and Forecasting Global Dynamics, reference EP/H02185X/1

Abstract

We present an international economic model which follows the input-output and computable general equilibrium (CGE) literatures in being based in large part on empirical observation. The model is a compromise between the mathematical elegance and linear simplicity of the first approach, and the theoretically rigorous, highly non-linear nature of the second.

Like both of these data-driven approaches, our new model is large-scale: each of up to around 200 countries is modelled with 35 economic sectors. This leads to challenges of both analysis and visualisation. We present some novel approaches to handling large-scale economic models which apply equally to input-output and CGE as they do to the model on which they are demonstrated here.

Additionally, the model presents a simple set of coefficient, which allow for it to be extended in ways which are explicable to non-specialists and policy-makers. We present the use of these coefficient sets to study four global human systems: trade, security, migration and international development. We make the case that these are too often modelled in the context of a single country or pair of countries. Our approach shows how placing these models in a wider context offers new, broader perspectives

Acknowledgements

Anyone who has shared an open-plan office with me over the course of preparing this thesis will know that I cannot think in any way other than "out loud". This implies that much of the work of this thesis was developed either in conversation with, or at least within earshot of, many of my closest colleagues at CASA. Particularly due a debt of gratitude are Thomas Evans, with whom I formulated the concept of this thesis over several very productive lunch times, and Rob Downes who was my sounding board and source of energy and inspiration for a very happy year sharing desk space. Thanks also to Adam Dennett for his boundless enthusiasm and for pointing me in the direction of doing a PhD in the first place and to Peter Baudains who listened patiently while I thought aloud at him. His advice was always sound.

I'm also very grateful to my friend and ally, Hannah Fry, who kept me sane throughout my PhD, via the medium of long walks and generally outstanding company. She also consistently told me what was misguided and, more often, what she loved about my work. Her wisdom and her sense of fun were both as invaluable as each other.

The unswerving faith and positivity shown to me by my supervisor Alan Wilson also deserves a special mention. I was more enthused about and more convinced of the work of this thesis after each supervisory meeting. That is a wonderful skill in a supervisor and it is much appreciated. Also, it was in conversation with Alan where many of the ideas in this thesis were first germinated.

A massive thanks to Anna Levy, Anna Leach, Helen Parnham and Pippa Hodges for providing the unending stream of motivational coffees I required to get me smiling into the start of so many working days. The Dream River cafe in North London has closed down since the completion of this PhD. I think it is clear to everyone that this cannot be a coincidence.

I'd also like to thank Anne Lund for acting suitably wowed by my enthusiastic sharing of probably quite inscrutable visualisations, and for listening to my presentations the night before I was due to deliver them and for saying that they were good even when they weren't.

Finally, of course, a huge thank you to Sheila and Mike Levy for their support (mostly emotional but occasionally also financial) and inspiration. It was they who first got me interested in economics, then gave me the idea to go to Germany for a Masters, then spurred me on to finish both that Masters and this PhD.

Contents

List of Figures

List of Tables

Part I

Introduction and literature review

Chapter 1

Motivation and approach

1.1 The need for global modelling

It has become popular in recent years to describe the world as an increasingly complex place. Commentators rush to claim that globalisation has made the world's economic system faster, more rapidly changing and more interdependent. The banking collapse of 2008 furthered the idea that an increasingly interconnected global system had become impossible even to understand, let alone manage or predict.

But despite this increased complexity, policy-makers still need answers to questions about the relationship between countries, their production technologies and their trading patterns. How would the global economy respond to slowing growth in China? What effect has the huge increase in migration in recent years had on the countries of arrival and departure? How does the current rapid development of some less-developed countries affect the wider economic system they are a part of?

These questions involve global human systems such as international trade, security, migration and international development. But in the literatures which study these phenomena, the truly global nature of the systems under study is not often enough acknowledged. The typical analysis of this type of question takes the form of a case study, selecting a particular country and looking at the effect of the phenomenon in question on the economy of that country. Regional extensions to this tend to be a set of such case studies, rather than an attempt at a unified modelling approach.

Beyond the single-country case study, economists wanting to show how the question at hand interacts with the wider economic system have traditionally started with a two-country model. This has been known to be unsatisfactory in understanding global human systems since at least the 1930s. As one early contributor to the global trade modelling literature puts it:

> Most of the important problems in international economics during the nineteenth century were discussed as though the world economy were divided into two regions, one region being the home country–usually England–and the other region being the "rest of the world." During the interwar period of the present century, this classical procedure came under heavy attack. (Metzler 1950, p. 351)

Economists sought to give a wider context to these inherently global human systems. Out of this desire was born the computable general equilibrium (CGE) tradition, which sought to combine national accounts in the form of input-output tables with the latest non-linear neoclassical economic theory. But despite its commitment to economic formalism, the rise of CGE modelling in the 1970s was accompanied by a schism among economists interested in modelling the economy at the country level.

By the time CGE reached its high-water mark with the IMF's multi-country Project Link (p. 46), the state of the art for theoretical economists was moving on from two-country models to those using the mathematics of infinitesimals. Here countries or sectors (or, sometimes, factors of production) were modelled as lying on a continuum between zero and one, with each occupying a negligible 'mass' of the whole (p. 74). This tendency for mainstream economists to count only in 1, 2 or infinity, which gave rise to the schism in economics, is born out of a desire for neat, closed-form mathematical solutions. But we would contend that this abstracts away from the precise heterogeneity and complexity which makes the global economy such an interesting subject of study. Today, theoretical economic modellers and CGE modellers rarely cite one another's work. The former seem to see something profane in the latter's focus on the often messy world of data, and the latter seem to be frustrated with the former's difficulty in translating their ideas to the real world.

1.2 A new data-driven approach

It seems clear that any attempt to model global human phenomena requires a large-scale model founded as far as possible on empirical observation. This model must contain a description of a country's economy which accurately reflects its nature, as distinct from the natures of the other countries of interest. We must be able to distinguish between developed and less-developed countries, between services economies and manufacturing economies, between an "Asian tiger" like Singapore and an OPEC member like Venezuela.

And the essentially interconnected nature of the systems under study in these questions requires us to understand not just the make-up of a country's economy, but crucially to be able to map out the dependencies between the parts of those economies. Additionally, the questions we have posed require us to understand the the international dependencies, which give rise to the second-order effects synonymous with globalisation. This leaves us with two possible multi-country data-driven methodologies: multi-regional input-output (MRIO) and CGE.

Both of these modelling traditions use input-output tables, and these are therefore at the heart of the work of this thesis. They make up the country models which are interconnected via a global trade network. Originating as a way to organise data at a sectoral level, they are a natural tool for the researcher looking to base a model as far as possible on real-world countries or regions. But MRIO and CGE lie at the extreme ends of a spectrum between simplicity and complexity in their use of input-output. MRIO has retained the mathematical elegance of Leontief's matrix inversion solution approach (p. 295) but has done so at the expense of model flexibility and the modelling of non-linearities. CGE has taken the path of modelling a wide range of economic phenomena, often via nested non-linearities, making the resulting models difficult to implement, and even more difficult to scrutinise and truly understand.

For example, the UK's Revenue and Customs CGE model for 2013 includes 11 non-linear equations, not including those associated with temporal dynamics. Of these, non-linear equations model such diverse phenomena as the production of finished products from intermediates, the consumer's choice between leisure and consumption and the substitutability of labour of different skill levels. The most recent iteration of a commonly used CGE model, GTAP (p. 64), lists 260 different variables with the behaviour of each country being modelled by 31 equations, not including those which govern the input-output part of the model.

We believe that a compromise between these two approaches is both possible and desirable. The great advantage of models built with mathematical elegance and linearity is that it is then easy to trace results back to those features of the model, either the data or the equations, which gave rise to the outputs. It also places empirical observation front and centre in its modelling approach: as little theory as possible is added in the move from data to model. When working with policy-makers, ease of comprehension and the traceability of results is genuinely important.

But CGE has the advantage of treating countries as separate units, and trade between countries as being fundamentally different to the exchange of intermediate goods within countries. The model we develop here has both advantages. It is a compromise between the easily understood MRIO, with its primacy of data and simple functional forms, and the flexible modelling of CGE, with its separable parts, each underpinned by its own economic theory.

To build this new model we use two relatively new sources of data. At the country level, a European Union project called the World Input-Output Database (WIOD) has recently provided input-output tables for 40 countries broken down into 35 sectors, for the years 1995 to 2011. A full description of WIOD is given in Appendix C on p. 302.

And at the international level, the UN has for a long time been gathering data about commodities trade and, to a lesser extent, services trade and these have recently been made publicly available via their website. These data sets are enormous, and managing them has been a not-inconsiderable part of the work of this thesis. A description of the commodities trade data set and some insight into the process of retrieving the entire set via the UN's web front-end can be found in Appendix D on p. 314. Each of these huge datasets provides important insights into the working of the global economy, and we will present separate analyses of both in this thesis.

By integrating into a unified framework the insights these two data sets provide, we can build a model of the global economy which is both simple enough to explain to a non-specialist, and rooted enough in empirical observations to make at least some claim to reflecting the state of the world as it really is. Although the relationships we propose between the elements of the model might seem like simplifications, we believe the benefits of clarity of exposition go some way to making up for the lack of realism. If we find a particular result to not be intuitively plausible (as we often will throughout this thesis) it is a straightforward matter to go back to the underlying data or the underlying relationships, and pinpoint where the problem lies. This allows a discussion with policy makers about issues in the data or assumptions, which might help the model improve incrementally.

1.3 Novel approaches to analysis and visualisation

As discussed, the model we present in this thesis has simple theoretical foundations, but is nevertheless large-scale, in the sense that it has thousand of moving parts. Each country consists of a set of internal intermediate goods flows and final demands, and can trade the output of each of its sectors with each other country. Visualising the results of such a large-scale model requires some innovation in analysis and visualisation, which we believe has been missing in the existing CGE and MRIO literatures.

The work of this thesis was done in the context of UCL's multidisciplinary "ENFOLD-ing" research project, and we hope some of this 'non-economics' influence has left a mark on some of our analysis and visualisation techniques. We have performed several analyses from the field of graph theory where we have treated the network of global trade as a graph with countries being nodes and trade relationships being edges (p. 193). We have also used clustering techniques from the field of unsupervised machine learning (p. 149). In studying consumption patterns and balanced growth we have used the mathematics of vector geometry (p. 203).

As well as these analytical techniques, visualisations such as the density plots on p. 111 and the co-occurrence matrix on p. 196 are new to the global modelling literature. These visualisations, and others throughout the thesis, are designed for the presentation of large quantities of data in a concise and visually appealing way. We think much of the economic modelling literature would benefit from modern visualisation techniques such as these.

These novel approaches, born out of working in a multidisciplinary environment, are as much a part of the contribution of thesis as the model we present, and more so than the particular results we find. We think of the model as being a vehicle through which such novel analyses can be conducted, and hence refer to it throughout as 'the demonstration model'. In several places in the thesis we will stress again this focus on the techniques and modes of analysis more than the particular results of this particular iteration of the model.

1.4 Structure of this thesis

In the next chapter of this first part we give an extensive overview of the history of global trade modelling and look in detail at some of the pivotal parts of the literature to which this thesis is a contribution.

The remainder of the thesis is then organised into two parts. In part II we explain in detail how the WIOD and COMTRADE datasets are brought together into a unified global model; how flows in the services sectors are estimated from the more sparsely-gathered UN services trade data; and what the estimated 200-country extension to the 40-country WIOD model looks like. With this the model is set up and ready to be analysed. We turn then to this analysis in part III on p. 157.

Part III comprises analysis in five broad topic areas. Firstly, we look at analyses of the model itself in chapter 6 on p. 157 and chapter 7 on p. 174. A major part of the work of this thesis stems from the fact that a global sector-by-sector model is, unavoidably, a large system. Changes to any part of the system, however small, lead to knock-on effects which can affect hundreds or thousands of flows in distant parts of the model. Summarising the effect of a particular change requires some thought and, as we have discussed, developing good and sometimes novel visualisation has been crucial to this. Chapter 6 and the other chapters in this section rely heavily on visualisations to help draw inference out of unmanageably large matrices of numbers.

The following three chapters (8, 9 and 10) then look at applications of our large-scale model to the study of global human systems. In each case, our goal is the same: to demonstrate how the use of a global model can bring new "whole system" insights to existing analyses from other disciplines. Thus we can, for example, explore which countries stand to benefit most from observed migration flows (chapter 8 on p. 206); how increased security threat levels between countries can benefit the economy of third parties (chapter 9 on p. 232); or which countries a selfish development aid donor might wish to favour (chapter 10 on p. 257).

More flexible than MRIO and simpler than CGE, our modelling environment allows for such second-order global questions to be asked, and for answers to be found which are based on real-world countries and on real-world observation. The analyses here are just a flavour of what is possible in such a context. We hope future users of such a model will find innovative, practical and useful analyses of their own, inspired by the examples given here.

Chapter 2

Global trade modelling literature

> *Statistical models of the working of the economy are not proposed as magic formulas which divulge all the secrets of the complex real world in a single equation. The statistical models attempt to provide as much information about future or other unknown phenomena as can be gleaned from the historical records of observable and measurable facts.*
>
> – Lawrence R. Klein, writing in 1947

To build an economic model, observations about the system under study must be brought together with theories defining how the various parts of the system affect one another when they change. In the case of modelling a system of national economies each connected to each via international trade, the economic systems can be summarised by the following questions:

- How does wealth/income influence expenditure?
- How does production technology define the translation of expenditure into quantities of goods and services?
- What determines the quantities of goods produced domestically and imported?
- What determines the quantities of imports sourced from each exporting country?
- To "close" the system, how do the acts of production and exporting affect wealth/income?

Each of these questions involves the "transmission" of one phenomenon (*e.g. income*) into anther (*e.g. expenditure*). The answers to each of these transmission questions come in pairs: which variables are involved in the transmission? In what manner and to what extent does the transmission occur?

This being *economic* modelling, the answer to the "which variables" question is usually "prices and other, non-price, effects." The second question, the form the transmission takes, is usually answered in a way driven by model simplicity and mathematical tractability. Relationships are almost universally specified as either linear (including log-linear) or a simple non-linear formulation such as constant elasticity of substitution (CES) or, in a far smaller number of cases, constant ratio of substitution (CRE). In either case, once the specification is decided upon, the task of modelling the form of the relationship between two variables comes down to estimating an *elasticity*, the extent to which a percentage change in one variable leads to a change in the other, also expressed as a percentage.

For two variables, x and y, the extent to which a change in x results in a change in y is described as the x-elasticity of y or, equivalently, the elasticity of y with respect to x. For small changes, the elasticity can be written as:

$$\frac{\Delta y/y}{\Delta x/x} \tag{2.1}$$

or, using the fact that $d(\log x)/dx = 1/x$, for small changes in x:

$$\frac{\Delta y/y}{\Delta x/x} \approx \frac{d\log y}{d\log x} \tag{2.2}$$

The practice of modelling the macroeconomy via elasticity estimations using data, either cross-sectional or time series, is considered by Dhaene and Barten (1989) to have started with a model of the Dutch economy by Tinbergen (1936). In a paper entitled "When it all began", Dhaene and Barten (1989) describe Tinbergen's model as the "first empirical model" and give a complete description of the model specification.

The Tinbergen model is conceptually advanced (although inevitably constrained by available data) and features many aspects of macroeconomic modelling which continue today. All the equations were estimated using data from 1923 to 1933 (an impressive feat in itself since the model predates the existence of the keeping of ordered national accounts. Tinbergen is said to have constructed several of the time series himself. The authors reproduce Tinbergen's model using more modern estimation techniques and conclude that "since the 1936 model, progress has no doubt been made; but less than might be thought" (Dhaene and Barten 1989, p. 203). It is therefore worth examining this model in some detail, as it shares features with many of the macroeconomic models which followed it.

2.1 Tinbergen's 1936 model

The aggregate econometric model grew out of a blend of several different streams of work...the mathematical economics stream springing from Walras...the work of Frisch and others in the theory of economic dynamics...the stream of work in statistical inference associated with Pearson and his successors...the development by King, Kuznets and others of numerical estimates of national income and expenditure...and the formulation of economic theories by Kahn and Keynes in the 'thirties. All of these streams are evident in Tinbergen's pioneering work.

– Carl F. Christ, *Aggregate Econometric Models*, 1956

Tinbergen's model (hereafter, TBG) is primarily a national model of the economy of the Netherlands, with the rest of the world (RoW) treated as a single entity. There are two social groups (labour and "others"), two types of good (consumer and investment), two manners of expenditure (consumption and savings), two measures of income (when earned and when received) and two stages of production (raw materials and finished goods.)

TBG introduces the notational convenience of expressing all quantities in terms of a percentage of the total annual wage bill, allowing all numbers to be expressed on a convenient scale (in TBG's particular example, all variables are expressed in units of 17.54 million guilders.)

The economy is assumed to have a trend growth rate, and all variables are expressed as fractional variations from a trend. The model can therefore be seen as a business cycle model, since fluctuations around a trend are all that is modelled.

TBG has 24 equations of which nine are accounting identities of the form:

total output = output for domestic consumption + output of export goods (TBG 5)

and

consumer expenditure = worker expenditure + expenditure by others (TBG 20)

The remaining equations then determine consumption, investment, exports, imports, total output, labour demand and capital income. These are determined by the exogenous variables which are the growth trend and every aspect of the RoW economy, i.e. all prices and quantities, except RoW imports from the Netherlands. Consumption is assumed to be a moving average of previous consumption levels along with a fixed marginal propensity to consume. Investment is a simple linear function of past prof-

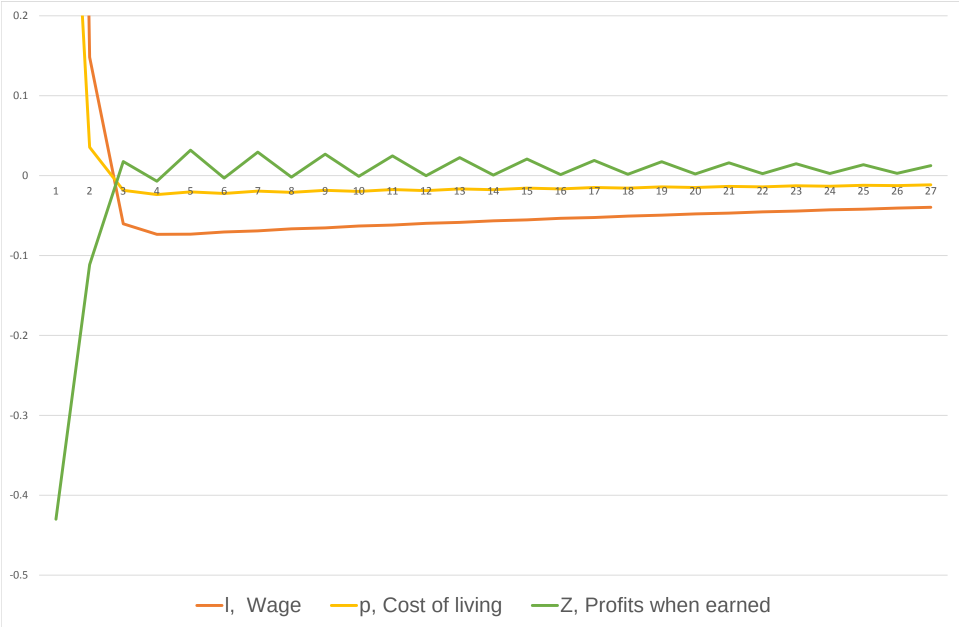


Figure 2.1: Impulse response of Tinbergen's 1936 model of the Dutch economy. Here, the model has been shocked with a one-off increase in the cost of living. Both the wage and the cost of living undershoot in response, and profits are net positive over the period of adjustment. We clearly see the model is well damped. For this simulation, we reproduced Tinbergen's model in Excel.

its and expected future profits with a trend which assumes investments grow over time. Rather than being an assumption, this is an empirical relationship, presumably a feature of the time period (1923-1933) over which the model was calibrated.

The level of exports depends on foreign and domestic prices (in short, the exchange rate) along with the volume of world exports, the price of which depends on world prices and input costs. Imports are represented as a choice between domestic and foreign goods depending on their relative prices. Total output is calculated as a fixed linear combination of imported finished goods (on which retailers make a mark-up), imported intermediate goods and a constant. The labour market is described as the level of employment in the production of consumer goods. Demand for this type of labour is correlated with output, and the price of labour (i.e. the wage rate) is a function of changes to cost of living and depends positively on total employment. This is typically Philips-type relationship. Finally non-labour income, profits and capital gains from capital ownership, is calculated as the difference between national income using factor costs and the wage bill.

The model is impressively complete, including all the main constituents of a modern understanding of the workings of the macroeconomy with the exception of two large aspects: the TBG model includes no monetary factors and no government or taxes. With 24 equations, several of which are difference equations, it would be helpful to see the dynamics of the resulting system in response to a simple shock. To that end, we have modelled the 24 equations as a simple set of Excel calculations. Figure 2.1 shows that model impulse response has complex, although damped, dynamics. The figure shows the model response to a one-off shock to the cost of living. The graph shows the difference between the shocked model and a model without that shock. After the initial cost-of-living spike, both the wage and the cost of living undershoot and return rather slowly to the non-shocked model. During this period, a positive profit is made with the net profit showing a saw-tooth response.

Numerically specifying a model of this complexity requires the estimation of a very great number of elasticities. Some of these, particularly those describing technological production processes, are based on information similar to a modern day input-output analysis, but all others are estimating using OLS from a time series of data covering 11 years. In stark contrast to a modern analysis, very little emphasis is placed on the standard errors of point estimates or indeed on the method used to calculate the many intercepts and elasticities. Instead the equations are merely presented ‘as-is’.

Nevertheless it is an impressive piece of work for its time, coming as it did during an era of paper records and hand calculations. In this context it is easy to see how questions of the robustness of the estimation methods and confidence levels would almost be seen as quibbles given the magnitude of the achievement.

The first major extension to TBG comes in the form of a multinational model by Polak (1939). Polak designed the econometrics to lend quantitative support (or otherwise) to two economic theories. Firstly, that improvements in the world economy lead to improvements domestically, due to increased global demand for the domestic exports. And secondly, that a weakening of the local currency boosts exports due to the fact that domestic production looks cheaper to world markets.

To these assumptions, Polak adds the further assumption that extent to which these two mechanisms affect the domestic economy depends on the openness to trade of the economy and on whether a country exports production or consumption goods. Openness is measured by the ratio of foreign trade to domestic income. The thinking is that a more open economy will be more sensitive to changes in world conditions. A country which exports production goods is then assumed to be more sensitive to global conditions and exchange rates than one which exports consumption goods.

Polak's econometrics does indeed back up these theories, with the exception of the USA which has a positive coefficient on its currency value with respect to its exports. This implies that the USA is able to export *more* when the US dollar becomes more expensive, presumably because of the status of the US dollar as a reserve currency. He then provides a model based explicitly on TBG, this being the only sufficiently detailed study to date. As he makes clear, the causal mechanism behind his econometric results "may be analysed in somewhat more detail for a country for which research in this direction has been carried out sufficiently, viz. Holland." (Polak 1939, p. 88)

As with TBG, all variables are measured as deviations from the average detrended values. Using the following relationship determined econometrically by TBG:

$$\text{Dutch exports} = \text{World exports} + 1.26(1.77\text{World prices} - \text{Dutch prices}) \tag{2.3}$$

and by showing econometrically that world exports correlate very closely with world production ($\rho = 1.0$), Polak shows that the first of his theses is proven, and domestic exports depend on "the world [business] cycle". Here, the world is taken to be eight "important" countries which together made up 55% of world trade in 1929. They are: USA, UK, France, Belgium, Czechoslovakia, Holland, Sweden, Norway and "other countries".

Polak then builds a simple model using the relationship found empirically:

$$\text{Dutch exports} = a_1\text{World output} - a_2\text{Exchange rate} \tag{2.4}$$

$$\text{Employment} = a_3\text{Import price} - a_4\text{Exchange rate} \tag{2.5}$$

making the assumption that the consumption good sector responds differently to the capital goods sector. The price of consumption goods and employment in the consumption good sector vary when the local currency devalues, but the capital goods sector is invariant with the exchange rate due to the cost of raw materials, which must be imported to produce such goods. Dutch exports are assumed to be a fixed proportion of total production, with the empirical multiplier found to be 1/1.56 (p. 92). It is therefore simple to show that business cycles in world output propagate through to cycles in domestic output via exports.

Polak goes on to say that this result holds for countries of "roughly the same economic structure" as the Netherlands (he includes the UK in this group) but that the USA requires a separate analysis, it being large enough to have endogenous business cycles. He models the world in such a way that only the US is allowed such a feature, an assumption he admits is "extreme, and not very realistic" (p. 94).

In rounding off their summary of the modelling environment in the 1930s, Dhaene and Barten (1989) conclude that the outbreak of war interrupted what had been a very fruitful period for such research:

> The next ten years were barren as far as the construction of models is concerned. However data bases were improved and methodological issues were tackled, so that when model building was taken up again the initial conditions were much more favourable than when Tinbergen was working (Dhaene and Barten 1989, p. 214).

This brings us to the post-war modelling work of the 1950s and '60s, which was summarised in 1967 in excellent detail in a review called simply "Models of World Trade" by World Bank economist Grant B. Taplin.[1] The next section gives a summary of this review.

[1] This era of the literature is also reviewed by the equally exhaustive, but somewhat earlier, review by Prais (1962). Since this review is earlier, Taplin's will be our main source here, but Prais will be called upon whenever his critiques are pithier.

2.2 Taplin's "Models of World Trade"

Taplin's World Bank review begins with a useful definition of the goal of trade modelling: "The ultimate goal of our work is to be able to forecast the level and the commodity composition of trade between any two regions" (Taplin 1967, p. 435). He contrasts this global approach to trade forecasting with the way domestic economic models of the time treat the foreign sector

> ...either as autonomous and predetermined or as a function of domestic factors exclusively. The international economy, however, is a complex network of interrelated trade flows, capital movements, and payments settlements. It is a system in which domestically induced changes in one country's income, prices, and other economic forces affect economic activity in other countries, which in turn transmit the changes on to each other and to the country of origin (p. 433).

This description of a "complex network of interrelated trade flows" shows that the recent fields of complexity and networks have a long intellectual heritage and that it was data, computing and theoretical constraints which led early modellers towards simplistic, mechanistic system representations not a belief that this was how the system really worked.

Taplin goes on to identify a divide in the global modelling literature which persists to some degree to this day. Descriptions of the global economy are split into those using a "model" from those using an "import-export matrix". The trade matrix versions covered by Taplin are:

1. The League of Nations's *Network of world trade* (League of Nations 1942)
2. Woolley (1966)
3. Beckerman (1956)

The descriptions which use a model are:

1. Tinbergen (1962)/Pöyhönen (1963)
2. Linnemann (1966)
3. Waelbroeck (1967)

And finally, there is a section of papers grouped together under the rubric of "short-run transmission mechanisms," in other words, within-country explanations for the levels and patterns of trade. These are transmission mechanisms in the sense that they seek to explain how fundamental changes at a local economic level are transmitted to international effects. Or, as Taplin himself describes it, using a quote from Neisser and Modigliani (1953), "the main relationships between the level of domestic economic activities in the various countries and their international transactions" (p. 5). This last group of papers consists of:

1. Metzler (1950)
2. Neisser and Modigliani (1953)
3. Polak (1954)
4. Rhomberg (1966a)

We will now give a short overview of each of these papers in turn.[2] The League of Nations trade matrix study is described by Taplin as:

> perhaps the first important study to attempt to fill in the import-export matrix... It's purpose was "to describe the pattern of [the] network as it presented itself before the...war and to consider how far that pattern was determined by the natural distribution of resources and how far by other factors of a more ephemeral or less unalterable character" (Taplin 1967, p. 436).

The study included exports (measured f.o.b) and imports (measure c.i.f.) for 17 regions / countries using trade data for 1928, 1935 and 1938.[3] Beyond these basic points, Taplin does not cover the League of Nations data tables in any more detail. But it is clear that this study was a forerunner of the United Nations' very extensive trade data-gathering exercise which culminated in the UN COMTRADE database, a key ingredient in the model of this thesis.

The book of Woolley (1966) gathers together trade data, primarily from the IMF's "Balance of payments yearbook" and presents the result in matrix form. The aims of the study are to present the statistics, describe the trade flows between nations and to do some basic analysis on the data presented.

---

[2] With the exception of Waelbroeck (1967) who attempts to predict future instances of the trade shares matrix using purely time series methods. Beyond Marwah (1976) we can find no other serious work which attempts to replicate this approach. Instead, future trade shares were forecast using fully specified models of the global economy, such as those described in the section on CGE modelling below.

[3] The Free On Board (f.o.b.) value of an export good is the value of the good at the point it reaches the exporter's national border. Thus it includes no transportation or insurance costs. It can be thought of as the "click and collect" of the trade world. The Cost, Insurance and Freight (c.i.f., formerly known as "Carriage, Insurance and Freight") value includes everything required to get the product to the *importer's* national border. It can be thought of as "including postage and packing."

Each entry in Woolley's nation-to-nation trade matrix has two recorded items: the trade flow figure as reported by the exporting country, and the same figure as reported by the importing country. Despite some of the discrepancies being large, a review of Woolley's book reports that "on the whole the pairs of corresponding records compiled from statements of different reporting countries agree surprisingly well." (Rhomberg 1966b, p. 1326). Discrepancies between exporter- and importer-reported figures is one which is no nearer to being resolved today that it was in the 1960s. Scholars may rely on the exporter's figure where available, falling back on the importer's reporting where the exporter data is missing. Alternatively an average of the two figures can be used. In this thesis, we use the former method.

The reviewer goes on to comment that the complications involved in gathering, recording and reproducing the data meant that Woolley had focused primarily on the first of his goals, that of presenting statistics, at the expense of doing any analysis. Finally, the reviewer summarises the book saying that due to the time period covered by the data being 1950-54, the study was already "chiefly of historical interest" by the time it was published. This is a perennial problem with data-driven studies such as Woolley's: when the data has finally been gathered, stored, analysed and presented it often seems rather out of date. This is a problem with the present thesis too, and is particularly acute when it comes to the use of input-output tables, which are compiled quite rarely, as they take a huge amount of quantitative research effort to produce.

The final study presented by Taplin in the category of import-export matrix is a mid-1950s study of "the multiplier" in world trade, the export response of a country to an increase in import demand (Beckerman 1956). This work builds a model of the network of world trade based on ten countries/regions (confusingly referred to by the author as "sectors", not be confused with ordinary sectors of the economy used in a modern context). Beckerman makes the link explicitly with the matrix techniques from input-output analysis, describing his work as "an experiment in the application to world trade of the techniques of matrix operation which have become familiar in economics through their application to input-output analysis." Beckerman's work can therefore be seen as a forerunner of much of the multi-region input-output work which contains the insight that parts of an economy are interdependent in reciprocal ways, with growth in one part leading to growth in many, if not all, others. It is a logical extension of this thinking to see the global economy as working in much the same way–a rather modern, globalised view of the way the world works.

The assumption from input-output that inputs are required proportionally to outputs is here translated into an assumption that imports are required in fixed proportion to exports. Beckerman refers to this fixed ratio as the "propensity to import."[4] This is therefore a purely global economy, with an almost non-existent model for the internal workings of a country's economy. The more a country wishes to export, the more it must import, in a proportional way.

The rather severe assumption that imports are a fixed proportion of exports is glossed casually over by reference to the short-term nature of the assumption, and to government manipulation which attempts to force the relationship to hold:

> It seems reasonable to suppose...that the factors internal to any given country which determine the impact on its imports...of a change in its exports are not likely to vary considerably during a period of about one year because the most important of these determining factors is the pattern of propensities to consume and invest within the economy... In addition, it has been the practice of several governments to "help" the normal multiplier process in a manner aimed at ensuring that the ratio between marginal changes in imports and exports is kept not only stable but stable at a value close to unity (Beckerman 1956, p. 241).

This model of simple responses of imports to exports is completed by a description of the USA in which that country is treated differently to all others.[5] Here, in contrast to the other countries, imports are completely decoupled from exports. Thus, in model terms, flows to the USA go nowhere, and flows from the USA come from nowhere. The USA thus acts as a "source" and a "sink" of global production and exports respectively, much in the way that the present thesis uses a RoW entity to resolve global trade imbalances (*see section 3.6 on p. 106*). The multiplier of a particular sector is defined as "the ratio of the final change in [the exporter's] total exports to the autonomous change in total US imports." Beckerman's model is therefore an import-driven global model with the USA playing the role of exogenous import-demand generator. In being import-driven, the model is similar to that of this thesis, albeit without an internal country representation.

Taplin devotes little time to Beckerman's model, asking why "propensities to import" should remain fixed, and even why imports should depend on exports at all. Indeed, the whole data-driven approach represented by the League of Nations, Woolley and Beckerman is dismissed by Taplin who says that although they give important insights into the structure of the global economy, they "do not present a way of rigorously testing hypotheses, measuring the impact of various determining trade flows, or forecasting

---

4 Compare this with the "import ratios" of this thesis, in section 3.3.2 on p. 89.

5 The other countries/regions are Canada, "dollar Latin America", Sterling countries of OEEC, continental western Europe, overseas territories of the above, overseas Sterling area, non-dollar Latin America, eastern Europe and "others."

future levels of both total and country trade" (Taplin 1967, p. 438). In this, Taplin represents an early example of a long-running tradition in the literature to be suspicious of data-driven approaches which eschew theory in favour of observation and minimum-possible modelling assumptions. A form of this critique can be applied to our model too.

Taplin therefore turns to models of global trade which seek not just to describe trade flows between countries, but explain and predict them too. The models Taplin discusses are all variants of a gravity model, to which the next section is dedicated.

2.3 Gravity models of world trade

In the early 1960s it became common to model geographic flows between two economies of different size in a manner similar to that used to measure gravitational interaction between two bodies of different size. In physics, two bodies exert an equal and opposite force on one another, with the force being proportional to the size of the two bodies, and inversely proportional to (the square of) the distance between them. Thus large bodies tend to exert a greater force, as do bodies which are closer together. The analogy in the trade context is that large economies tend to produce greater trade flows, and that economies which are closer together will tend to trade more than those which are far apart. Thus the gravitational force of the gravitation equation can be thought of as being analogous to the quantity of traded goods between two countries.

It is widely understood that this physical analogue was invented in 1962/3 simultaneously by a Dutch researcher and by a Finnish one. What makes this story particularly entertaining is that the Dutch researcher, Jan Tinbergen, appears to have invented the formulation in a moment of spontaneous inspiration in an appendix to an otherwise unrelated book, a utopian attempt to remodel the world economic order (Tinbergen 1962, pp. 262-293). But, contrary to this often cited origin story, Tinbergen in fact mentions earlier in the book that the choice of size and separation as the main variables determining trade flows was made "on the basis of earlier and more detailed empirical studies conducted by Mr. G. J. Ayelts Averink, at the Netherlands Economic Institute" (p. 60). No further detail is given and Appendix 6, which makes up the body of Tinbergen's contribution to gravity modelling, and is presumably the only part of Tinbergen's work which many scholars read, does not mention this researcher anywhere.[6]

An unfortunate postscript to the story of simultaneous invention is that the Finnish author, Pentti Pöyhönen, was aware of Tinbergen's formulation as having been published in July 1962, predating his article, saying that "one and the same theory seems to have been elaborated simultaneously but independently at two different research centres." (Pöyhönen 1963). A note from the editorial department just above this reads: "We received the present paper already on the 3rd of November 1961. Because of lack of space it could be published only just now." Due to this publishing decision, Tinber-

[6] There is somewhat little trace of GJ Aeyelts Averink in the online literature databases, but he seems to have worked in the fields of marketing and management. See, for example, Panne and Averink (1961). Additionally, Aeyelts Averink and Tinbergen are mentioned together in a list of staff working at the Econometric Institute of the Netherlands School of Economics in Rotterdam (Theil 1961), supporting the idea that one might have learned from the other. But there is no trace in Aeyelts Averink's published work of a nascent gravity formulation.

gen is now widely credited as having invented the gravity model of global trade and Pöyhönen is an altogether forgotten name.[7]

Despite this (possibly erroneous) crediting of Tinbergen with the invention of the gravity model, the use of such formulations was in fact already widespread by the 1960s, but just not in economics. It was common in the similar but seemingly disconnected field of economic geography where it was described under the rubric of "spatial interaction," the geographers calling global trade modelling "commodity flow analysis". A 1970 paper on commodity flow analysis mentions a large number of gravity models of international trade which far predate the models of Tinbergen/Pöyhönen (Smith 1970). Indeed, without even citing Tinbergen, Smith gives the following description of a strand of literature which perfectly describes what Tinbergen sets out to do:

> A common objective in many commodity flow studies is to make diagnostic statements about the magnitude of flows. Such statements usually take the form of "more than" or "less than" might be expected under certain conditions. The observed flows are arrayed against "predicted" flows over the given routes, and inferences are drawn about the relative magnitude of the actual flows (Smith 1970, p. 406).

The same author claims that a hypothetical review of the literature on gravity and trade flows would be "extensive." Furthermore, Carrothers (1956) gives a "historical review" of the use of gravity formulations in human interaction which goes all the way back to the first half of the 19th century. This latter review goes on to describe the evolution of the gravity model in studying migration in the mid-1920s (Young 1924) and a "law of retail gravitation" of 1931 (Reilly 1931). So the use of gravity-like equations to study flows between sets of origins and destinations far predates economists' use of the 1960s, but Tinbergen's credit as the originator of the model have nevertheless led to his work receiving a very great deal of attention, and it is for this reason that the original study deserves looking at in detail.

2.3.1 Tinbergen

The purpose of Tinbergen (1962) in developing a model to "explain" the pattern of trade is much more clearly defined that that of Pöyhönen (1963), and it is explicitly political in nature. The book in which "An Analysis of World Trade Flows" appears as an appendix is concerned with the way in which historical discrimination in the global

[7] As far as who deserves credit for thinking of the mathematical formulation as being similar to the gravitation equation of physics, Pöyhönen mentions the physical analogue specifically, noticing "the close connection between our estimator function and the gravitational problem of two bodies" (p. 99) but Tinbergen does not talk about gravity or physical analogues at all.

system has led to some countries to be over-represented in trade volume terms, and others to be unfairly overlooked. Tinbergen's model is therefore a first attempt, which he himself describes as "very simple", at determining "the normal or standard pattern of international trade that would prevail in the absence of discriminating trade impediments" (p. 262). In other words, Tinbergen wanted a matrix of world trade flows that one would 'expect' to see given only fundamental facts about the economies involved, in order to see where the observed trade flows most differed from this 'expected' norm. In doing this he hoped to uncover preferential treatments and discriminations which he saw as being "obstacles to the optimum flow of international trade" (p. 262). Tinbergen offers no justification as to why his gravity-like formulation should be expected to lead to "optimum" flows of trade, and should therefore be seen as a proof of concept, demonstrating the kinds of techniques that might be applied to the question of "fairness" in global trade. This is a subtlety which we have never seen commented upon in the very voluminous gravity literature which followed.

Tinbergen's first specification, described as the "simplest form," is given as:

$$E_{ij} = \alpha_0 Y_i^{\alpha_1} Y_j^{\alpha_2} D_{ij}^{\alpha_3} \tag{2.6}$$

where E is the volume of exports, Y are Gross National Products and D is the geographic distance.[8] In this original formulation, geographic distance was taken to "correspond roughly" to transportation costs, and are therefore to be taken as a proxy for the latter.[9] As Tinbergen points out, this formulation implies a constant elasticity of trade with respect to each of these variables.

The α's were estimated using data from 42 countries (thus 42×41 observations) in 1959, which covered around 70 per cent of total world trade in that year. The results were $\alpha_1 \approx 1.0$, $\alpha_2 \approx 0.9$, $\alpha_3 \approx -0.9$ with an R^2 of around 0.8. These results were found to be reasonably robust to the addition of extra dummy variables for shared land border and for various political "preferences" such as Commonwealth membership and

[8] Tinbergen goes on to add several more explanatory variables to the specification, although these are all dummy variables attempting to account for "preferential treatment" among various groups of countries. In doing this, Tinbergen blurs the line between attempting to explain deviations from some "normal" pattern of trade and attempting to maximise the R^2 of his calibrating regression. In doing this, he is the first in a very long line of researchers who fail to make this distinction.

[9] Beyond the very first papers on the topic, it quickly became rare to use this "simplest form" gravity equation with only distance and GDP. A modern exception to this rule comes from two researchers at the Tinbergen Institute in Amsterdam. Bun and Klaassen (2002) estimate this simplest form, but do so using a panel data estimation, with lagged values of both the trade variable and GDP. One unusual side-effect of this method of estimation is that distance, such a crucial variable in all early instances of the gravity model, is subsumed into the bilateral fixed effects—the non-changing aspects of each ij pair.

Benelux membership.[10] He thus concludes that exports are "almost proportional to the GNP of the exporting and the importing country" but that the small difference in the two coefficients appears robust to many different specifications and data sets. There is therefore held to be strong evidence for an effect whereby "large countries (in terms of GNP) always export more to small countries than they import from them; this leads to a positive balance of trade for the bigger countries and a negative trade balance for the smaller ones" (p. 289). Tinbergen also finds some evidence that a more diversified export offering leads to a greater export volume. This finding, barely a footnote to Tinbergen's primary discussion, prefigures the far later, and far more famous, result of Hidalgo and Hausmann (2009) that more advanced economies benefit by producing a greater range of products.

Tinbergen then goes on to discuss deviations from the expected flow values given these parameters, with particular focus given to negative deviations indicating "discriminatory trade impediments" (p. 292). He emphasises that his results overturn the standard (at the time) assumption that industrial countries trade more with one another than do non-industrial ones, explaining that this discrepancy is fully accounted for by the size of the economies in question. He finds that the only countries in his sample which could properly be called "trade victims" are Pakistan and India (p. 66). Other countries show, for example, "subnormal imports" but not subnormal exports. He says that these imbalanced trade deficiencies may be due to those countries' "own autonomous restrictions." He draws the conclusion from this that the then popular focus on boosting trade in developing countries, rather than the giving of aid, may be inappropriate, since those countries' trade patterns are broadly what one would expect given the size of their economies: "...it follows that trade policies should not be overemphasized as an instrument in development; thus, we do not think the slogan 'Trade, not aid' a proper guide to action" (Tinbergen 1962, p. 66).

2.3.2 Pöyhönen

In contrast to Tinbergen, the paper of Pöyhönen (1963) is concise, but it is similarly rather light on commentary about how the specifics of his model were selected. However the simplicity of his exposition and of his approach are admirable.

He begins with a matrix approach, describing "the exchange of goods" (p. 94) with rows representing exporters and columns representing importers. He calls this trade matrix A. He then posits a general matrix function which would use matrices of other information, beyond that in A, to estimate A, in the sense of reducing the mean squared error between the elements of A and those of the estimated matrix. In other words, Pöy-

[10] Benelux is short-hand for Belgium, Netherlands and Luxembourg.

hönen is attempting to 'explain' the observed trade matrix using non-trade covariates. Using a matrix in this way leads naturally to a comparison with input-output analysis. He points out the link to, and difference with, this widely-understood technique:

> The simplest possible model would result from an exclusive use of matrix *A* and a number of parameters, so that the explanatory function...would be purely structural as is the case with the input-output analysis. The amount of additional information is relatively small in comparison with the information contained in the original matrix... An explanatory model in the true sense of teh word can be spoken of only when information in the form of other matrices is utilized in addition to the basic matrix *A* (Pöyhönen 1963, pp. 94,95).

In other words, input-output analysis is purely structural or, as we might say today, "data driven," whereas the model Pöyhönen is introducing is an attempt at explanation. To this end, he introduces two new matrices of information: a diagonal matrix of national incomes and a square matrix of transportation distances between the major ports of each country.[11] No explanation is offered for why these two information matrices are selected, nor is any given for the choice of functional form, other than that it is "quite general". The equation given (eq. (7), p. 95) is:

$$a'_{ij} = cc_i c_j \frac{e_{ii}^{\alpha} e_{jj}^{\beta}}{(1+\gamma r_{ij})^{\delta}} \tag{2.7}$$

where the c's are constants, the e's are national incomes and r is the transportation distance.

Pöyhönen finally goes on to estimate this equation using logs, and finds that α and β are both close to 0.5 and that δ is close to 2. He concludes by stating that "regional studies have frequently shown that economic activity is governed by rules of which the analogies are to be found in natural science" and goes to discuss "the close connection between our estimator function and the gravitational problem of two bodies" (p.99).

The specification and the physics analogy, if not the precise parameter values, are very similar to much of the gravity modelling which followed Pöyhönen, but this work is rarely cited and the man himself is now largely forgotten.

2.3.3 Other early gravity models

An early application of the gravity formulation of Tinbergen/Pöyhönen was an attempt by Waelbroeck (1964) to estimate the trade impact of the European Common Market,

[11] The diagonal elements of this distance matrix are, of course, set to zero.

created five years earlier by the treaty of Rome, along with the European Economic Community (EEC).

Waelbroeck finds a neat application of the gravity model to assess observed trade flows against those that would be "expected" given the gravity formulation, after the creation of the Common Market. He estimates equation (2.7) for 1958, before the creation of the EEC, then assumes that the constants c, c_i and c_j are unchanged over time. This then allows him to extrapolate the 1958 trade matrix forward to 1962, after the creation of the EEC, to compare the hypothetical trade levels with the observed ones. By doing this Waelbroeck finds that intra-EEC trade had increased "considerably more than the Finnish model would have led us to expect" and that there was no evidence of trade diversion from North America or from other European, non-EEC, countries.

In a review of related literature, Balassa (1967) describes previous general-equilibrium studies as suffering from a confounding of the general improvement in Europe's economic "weight", diverting trade demand from elsewhere, with the genuinely trade-increasing effects of the Common Market. He describes works which seeks to overcome this problem by analysing drift in the trade share matrix (which we will look at in detail later in this chapter) but argues that "by proceeding in a piecemeal fashion, it does not provide fully consistent results." (p. 2). He then describes how Waelbroeck seeks to distinguish between the trade-creating and trade-diverting effects by applying the gravity work of Tinbergen and Pöyhönen.[12]

Belassa does criticise this work however, for the fact that the income elasticities (α and β in equation (2.7)) are calculated "in a cross-section analysis of *all* trading countries...[whereas] elasticities are generally higher in the industrial countries." (p. 4). He cites internationalisation as a cause for higher income elasticities in industrial countries and protectionism as a cause for the opposite effect in less developed countries. He further warns that using aggregate trade levels may hide changes in different commodities in different directions. He completes his analysis by estimating income elasticities for each of eight categories of goods, both before and after the creation of the EEC.[13] This allows a quite detailed analysis of the 'before' period (1953-59) and the 'after' period (1959-1965) in the European trading economy.

This focus on calculating income elasticities of import demand is described by Aitken (1973) as creating a distinct strand of the literature in comparison to those papers which

[12] Balassa does not mention "gravity" in his description of their work, describing it simply as "regression analysis, with gross national products and geographical distance as the principal determining variables." (p. 3). Interestingly, in light of the above discussion about the origins of this model, he gives an overview of "the Finnish study" and not one of Tinbergen.

[13] The eight commodity categories are, broadly, food and drink, raw materials, fuels, chemicals, machinery, transport equipment, and "Other manufactured goods."

use "income as an independent variable in a trade-estimating model." Thus we see a point of separation between those studies using trade flows as a starting point to estimate import elasticities for broader modelling purposes, and those estimating trade flows as an end in itself. This is a crucial separation, which remains to some degree throughout the literature on trade.

Another early extension to the Tinbergen/Pöyhönen model is given by Linnemann (1966) who distinguishes between the 'size' of a country as defined by its economic heft and a country's size in population terms.[14] He thus allows the gravity formulation to give different treatment to populous but poor countries, which are likely to have more trade simply on the basis of the size of their population. His empirical efforts at explaining trade flows come in the context of theoretical work by "neo-classical international trade theorists" whose comparative advantage and factor endowment theories had "proved particularly disappointing in providing concrete guidelines for empirical work." (Ball 1967).

Excluding communist countries and "many minor trading nations" (Sawyer 1967), Linnemann has a matrix of data for 80 countries averaged over the years 1958-1960.[15] He adds both exporter and importer populations to the original specification, along with his own set of "preferential trade" dummy variables to represent the various trading blocs in existence at the time, such as the British Commonwealth. In order to improve the goodness of fit, a "commodity composition" variable is added which attempts to adjust for the extent to which the commodity make-up of an exporters offering matches the make-up requirements of an importer. The idea is that countries are more likely to trade with one another if what the exporter has to offer matches what the importer wants to buy. One reviewer concludes by saying that:

> ...in cross-section analysis of this type a large proportion of the variance of trade flows [are] left unexplained...relating to the effect of financial arrangments, aid and capital markets... we are still far from a model of great predictive power that would assist us in predicting the pattern of trade (Ball 1967, p. 368).

Despite this, Linnemann's work led to a flurry of studies which adjusted the formulation and search of an improved goodness and fit, and an attempt to deal with some "surprising" empirical results.[16]

---

[14] An interesting overview of Linnemann's career at the Netherlands Economic Institute, and his close working relationship with Jan Tinbergen, is given by de Hoogh (1994).

[15] Ball (1967), Sawyer (1967) and Smith (1970) are used extensively in this section's description of Linnemann. Each is (or, in the case of Smith, contains, a contemporary review of Linnemann's book.)

[16] As part of an extensive review of the literature on commodity flow studies, Smith (1970, p. 408) praises Linnemann's work saying "the important role of the distance and 'mass' variables should not be overlooked, although the negative exponents for population...are surprising."

Yeates (1969) attempts to bring the developments in trade model now familiar to economists to the attention of geographers.[17] He produced a short "note" which followed a paper presented at a conference a year earlier, concerned with "demonstrating that the volume of trade between countries can be explained by a modification of the well-known gravity model." To this end, he proposes an simple version of the work of Linnemann and Beckerman (1956, reviewed above), including what he calls "trade groups." He also employs a slight technical advancement of calculating great circle distances between country capitals.

By calculating a goodness of fit measure for each of a set of "arbitrarily chosen countries," Yeates finds that the UK and France are unusual in having $R^2 < 50\%$ and point estimates that look different to those of the other countries. He adds a dummy variable to former colonies of the UK and of France and finds a greatly increased goodness of fit. We note at this point that although Yeates states that Beckerman and Linnemann "have not...incorporated [distance and economic wealth] along with trade groups into a model relating to interaction in general," it seems to us that Linnemann did exactly this with his model, and that Yeates is not adding anything new here. But perhaps the new audience (being geographers, not economists) justifies this short paper's inclusion in this review.

The economic geographer Robert H. T. Smith starts his work on commodity flow analysis (Smith 1970) with a complaint that the geography version of trade analysis suffered from a "conceptual poverty," with many geographers being content to visualise flows, rather than to explain them. After an overview of spatial interaction modelling, Smith gives a list of those researchers who were interested in the effect of distance on trade volumes, before outlining the gravity model of trade, which he attributes to Linneman, claiming that the distance and population variables are "the only two variables with any general applicability...most others seem to be peculiar to a given situation." What follows is a review of linear programming methods to establish the most efficient routes for transportation of goods, none of which relate to international flows, and early graph theory methods of analysing the flow matrix, such as that of Nystuen and Dacey (1961), whose work inspired the network analysis of Levy, Oléron Evans, and Wilson (2016). But none of these examples attempts to look at global trade, all being specific regional studies.

A large strand of trade literature centres around estimating income elasticities of trade. In other words, how much more a country will import if its income grows by 1%.

[17] Yeates cites two studies which look at commodity flows using the tools of geography, rather than economics, *viz.* spatial interaction. Both these studies, Reed (1967) and a work in the book by Berry (1966) are on the specific subject of the Indian economy and are not in the scope of this review.

Although this can be viewed as a more narrow question that that approached by gravity modelling, since only imports are measured and not bilateral flows, it nevertheless seems to us that at least Houthakker and Magee (1969) are worth mentioning here.[18] We look in more detail at the literature around elasticities in a later section on p. 69.

In Houthakker and Magee's paper, which cites none of the gravity papers, but which *does* cite many of the papers in the next section on "transition mechanisms," they employ a multiplicative model not dissimilar to the gravity formulation.[19] They seek to explain the import level, M_{it}, in terms of the importer's income, Y_{it}, and a ratio of an import price index, PM_{it}, and an overall price index, P_{it}:

$$M_{it} = \alpha_{0i} Y_{it}^{\alpha_{1i}} \left(\frac{PM_{it}}{P_{it}} \right)^{\alpha_{i2}}$$

The authors take logs and estimate the α's, along with the parameters for a similar export equation, for fifteen industrial countries between 1951 and 1966. The focus of their paper is on questions around growth and balance of payments, which is a different focus to the literature reviewed here on explaining the levels of global trade. This might explain the lack of 'cross-pollination' between this paper, and the literature which stems from it, and the gravity literature which, on the surface, seems to be attempting to answer the same question.[20]

In this section we have reviewed that various early attempts at explaining the quantities and, to a lesser extent, the product mix, of international trade, using macroeconomic observables about origin and destination countries. In none of these studies is the internal condition of the economies taken into consideration. We now look at two more modern studies, both extremely influential, which brought gravity models into what we might call "economics respectability" by arriving at the formulation via ideas from neoclassical consumer/producer theory. These works were instrumental in bringing to the gravity equation, as one of the authors puts it, "gravitas."

[18] Paul Krugman later described the work favourably, saying that it "remains a benchmark for comparative estimation of trade equations across a large number of countries. Their main conclusion was that there were large differences among countries in their relative income elasticities — specifically, that Japan faced the highly favorable combination of a high income elasticity of demand for its exports and a low income elasticity of import demand, while the U.S. and the U.K. faced the reverse" (Krugman 1989, p. 1034).

[19] Transmission mechanisms, as we shall see in the next section, describe how imports are related to exports, thus how changes in exports are "transmitted" through a country model to imports.

[20] In a paper we cover on p. 69 of this review, Goldstein and Khan (1985) give a very detailed overview of Houthakker and Magee (1969) and criticise their paper "and indeed of all income-based trade models" (p. 1084) in three ways. In short, the lack of a limited supply capacity, the lack of differentiation between secular and cyclical (i.e. systemic vs. short-term) income movements, and that the models are "too partial-equilibrium." It is though worth noting that Goldstein and Kahn do not address any of these criticisms in their own models.

2.3.4 Neoclassical gravity models

Based on work in the late 1970s by Boston economist James Anderson, Anderson and Wincoop (2003) (hereafter AvW) show how Armington's differentiation of goods by origin, and his constant elasticity of substitution (CES) functional form for consumers' preferences, can lead theoretically to a formulation which shares many of the characteristics of the gravity equation.[21] This paper is almost ubiquitously cited in the economics trade literature which follows, and is therefore worth looking at in a little detail.[22]

They begin with a set of countries or regions which produce only one good, earn income only through the production and export of that good, and distribute all income to a representative consumer. This allows them to use good and region interchangeably, which they do throughout their exposition. The representative consumer of country j maximises a variety-loving, CES utility function:

$$U\left(\{c_{ij}\}\right) = \left(\sum_i \beta_i^{-\rho} c_{ij}^{\rho}\right)^{1/\rho} \qquad \text{(AvW eq. 4)}$$

subject to the budget constraint of the country's overall income, y (assumed, recall, to be the same as its total production/export):

$$\sum_i p_{ij} c_{ij} = y_j$$

The utility function has the important feature that country j's utility depends on its consumption of goods from *every* country i, with the extent of variety loving being determined by the set of "distribution parameters," β_i.

Each region faces different trade costs with each other region, which scales the exporter's price of their good by a factor, t_{ij}. Thus, each bilateral trading partnership is associated with a different price $p_{ij} = p_i t_{ij}$. By assuming that prices are determined such that markets clear, and that trade costs are symmetric ($t_{ij} = t_{ji}$) they are able to find an equation for prices *and* an equation for the value of each ij bilateral trade flow.[23]

---

[21] See the section on Armington beginning on p. 43 for a discussion of CES functional forms, and Armington's ideas that products from different countries were not perfect substitutes for one another.

[22] For reasons of their subsequent analysis, AvW use $(\sigma - 1)/\sigma$ as their CES parameter. This is overly burdensome for our purposes, so we replace it here with ρ. Elsewhere, where required, we use other Greek letters to replace lengthy parameters whose particular form is not of interest to us here and in the remainder of this section. Furthermore, in our exposition of Helpman, Melitz and Rubinstein below, we replace their use of $\breve{p}$ with an unadorned p as the distinction is of no interest to us here.

[23] Notice that they are not solving for quantities, as the previous gravity literature did, but rather for value, which is quantity times price.

The effective price from the importers' perspective in AvW's formulation end up as a CES combination of every exporter's price and each of the importer's bilateral trade costs:

$$P_j = \left[\sum_i \left(\beta_i p_i t_{ij} \right)^{\lambda} \right]^{1/\lambda}$$

This price index responds to changes in any of the distribution parameters, any of the exporters' prices, and any bilateral trade cost. Because any of these three components represents a dissuasion to trade, and because each one depends on variables from all other countries, AvW describe the combination as a "multilateral resistance."

The multilateral resistances are then used in place of prices in their final gravity formulation which, they say, "has a much more useful interpretation" than the previous gravity literature which "pays no more than lip service to theoretical justification" (p. 174). Again, expressed in value, not quantity terms, the specification is:

$$x_{ij} = \frac{y_i y_j}{y^W} \left(\frac{t_{ij}}{P_i P_j} \right)^{\lambda} \qquad \text{(AvW eq. 13)}$$

where y^W is global income. By making the simplification that trade costs, t_{ij}, are a simple function of the distance between the two regions, AvW are able to estimate their equation and find they are able to estimate trade flows well using the USA and Canada as their regions. A simple and popular method of estimation based on Taylor-series approximations to the multilateral resistance term was provided by Baier and Bergstrand (2009).

A very well-known contribution to AvW's extension of the gravity equation acknowledges that, being a multiplicative model (i.e. log-linear), it has trouble with zero flows:

> By disregarding countries that do not trade with each other [we] give up important information contained in the data, and produce biased estimates as a result. We also argue that standard specifications of the gravity equation impose symmetry that is inconsistent with the data... we develop a theory that predicts positive as well as zero trade flows. (Helpman, Melitz, and Rubinstein 2008, p. 442)

The framework of Helpman, Melitz and Rubinstein (hereafter HMR) begins with a multi-country world each producing and consuming the same continuum of products, B_j. As with AvW above, the utility function is CES in form, except with an infinite number of products, the sum becomes an integral, but otherwise has the usual CES form:

$$u_j = \left[\int_{l \in B_j} x_j(l)^\alpha dl \right]^{1/\alpha}$$

Demand, $x_j(l)$ is then specified as depending on income in j, along with the exporter's price, $p_j(l)$ and a CES-style price index similar to, although simplified from, AvW's:

$$P_j = \left[\int_{l \in B_j} p_j(l)^\rho \right]^{1/\rho}$$

With a comparatively detailed description of the firms doing the producing, using a monopolistic competition framework (which we look at in more detail on p. 72), a simple mark-up pricing model, and differing production efficiencies per firm, HMR are able to derive an expression for the operating profits of firms in each country. This then allows them to model the idea that only some, or even none, of a country's firms are willing to export to a particular other country. This is the innovation which allows them to model zero trade flows. It is also a non-symmetric relationship: just because country i's firms are productive enough to overcome the costs of trading with country j, it does not imply that country j's firms are productive enough to overcome these very same costs. This is an impressively complete theoretical analysis of trade, but the empirical section is stuck with the same limited data as all other empirical investigations of trade, and has to make several simplifications to be estimable.

Their empirical specification for trade volumes ends up including a fixed effect for the importer and the exporter, the geographical distance, and a variable w_{ij} which the authors claim "controls for the fraction of firms (possibly zero) that export from j to i" but which is really nothing more than a bilateral fixed effect. The real innovation, though, is that a first stage is run which predicts *which* countries choose to trade with one another. Their work is a theoretical *tour de force* but the empirical section is dealt with rather more hurriedly, and includes such oddities as the fact that a shared border is predicted to decrease the chances of two countries choosing one another as trading partners. Both this and the AvW paper discussed above are now very common in the trade literature, and more on their non-gravity precursors in given in the section on trade models with microeconomic foundations beginning on p. 72.

But we first give an overview of a much earlier strand of the literature which sought to use the theoretical description of the economy in explaining patterns of trade, by beginning with economic fundamentals such as investment and expenditure functions and building models of domestic demand that then result in the imports and exports that make up the global trade network.

2.4 Short-run transmission mechanisms

We now look at the papers which Taplin (1967) described as seeking to "establish the main relationships between the level of domestic economic activities in the various countries and their international transactions." Thus, instead of treating economies as "black boxes" with only macro-level observables such as population and GDP which, in some way, attract trade, these papers seek to explain something about the economies involved and how, for example, investment or income result in demand for and supply of internationally traded goods.

A very early and quite simple attempt at modelling a multi-country setup with an aim at explaining trade via domestic fundamentals is that of Metzler (1950). His model takes a familiar 2-country model, and expands it to *n* countries. The model traces the response of production and hence of trade to a "disturbance in the economic forces governing income" but in all other senses the model is completely static: prices, costs and exchange rates remain fix, and do not adjust in response to changes in trade patterns. Metzler justifies this static world by saying that his model is an attempt to model a future world in which international trade is able to happen in a frictionless way, as opposed to "the abnormal conditions prevailing today" of unbalanced trade, dollar shortages and exchange controls. The perfect trading world Metzler imagined never arrived (or, at best, existed only for a very short time up until the early 1970s) but his model is of interest in this present context, due to its innovative use of input-output tables at the heart of a more complex economic model.

Metzler proposes to use "the input-output tables developed by Leontief in the study of inter-industry relations" in a new way, with countries replacing industries, and inputs and outputs being replaced with imports and exports. A generalised import function describe how a country's import demand from a particular country depends on its income. These are then combined in an additive way to get a "total import function" relating domestic income to total imports from all exporting countries combined. Expenditure functions then relate total expenditure (of all kinds, domestic and foreign) also to income. These functions are then laid out in an input-output matrix with "receipts from sales by country *i*" down the rows, and "expenditures by country *j*" along the columns. This matrix is then solved in the normal matrix inversion way in response to, for example, a propensity to consume (a change to one of the expenditure functions.) In this way, an increased propensity to consume out of production can ripple through the global trade network creating an increased income, and hence in turn increased import demands, in all other countries in the model.

Metzler's modelling approach is very simple, but by decoupling the domestic equations from the international equilibrium, he shows how more complex, nonlinear, behaviour can be added between the solution of the matrix inverse problem and the response of the various equations that make up the row/column constraints. He proposes a "disequilibrium model" which describes what happens when, due to an exogenous change in one of the equations, supply no longer equals demand. This is done with a set of differential equations, one for each country $i \in (1...N)$, of the form:

$$\frac{dy_i}{dt} = k_i[u_i(y_i) - m_i(y_i) + m_{i1}(y_1) + ... + m_{in}(y_n) - y_i] \tag{2.8}$$

where the $u(\cdot)$ are the expenditure functions, the $m(\cdot)$ are the import functions, k are constants describing the speed of adjustment of output, and each function is subscripted with country number(s). Since Metzler is unwilling to make his model specific by giving functional forms to $u(\cdot)$ and $m(\cdot)$, he replaces each function with a linear approximation and checks for the stability of the system of equations as a whole.[24]

Following a long and technical discourse on proving stability of the system, Metzler goes on to measure the effects in his model of a permanently higher rate of investment in country 1 on the incomes of all countries. This is where the work becomes a genuine (if fairly simple) attempt to combine the complexity of a full n-country, possibly data driven, approach with the kind of economic modelling which usually takes place in much more abstract settings. Disappointingly for an exercise of this ambition (given the limited data storage and calculation technology available to Metzler in the early 1950s) he concludes that the n-country model gives general conclusions which are the same as the results of a two-country model. He justifies his studying of the n-country model by saying:

> Although our study of the n-country model has not taken us very far, it has, I fear, taken us about as far as we can expect to go without introducing actual numbers in place of our hypothetical propensities to import and to spend... Eventually, then, an import-export matrix, similar in many respects to Leontief's input-ouput matrix for a single country, must be developed for the world economy... With the improvement in statistics throughout the world since the end of the war, it is to be hoped that this gap in our knowledge will soon be filled (Metzler 1950, p. 354).

In contrast to Metzler's focus on rates of investment, the work of Neisser and Modigliani (1953) looks at how much international economic activity is required to

[24] By replacing a general function with a linear approximation, Metzler only need specify the direction of the relationship between the function and its arguments. He maintains that this is a worthy exercise since "stability of the linear approximation is obviously a necessary condition, although not always a sufficient condition, for stability of [the system]." (p. 337)

sustain a given level of domestic economic income. Their model ties imports to exports in a simple way. They begin by assuming that imports are a fixed proportion of output, which is exogenous. They then assume that a country's exports are a fixed share of the total imports of all other countries. In these greatly simplifying assumptions, their model is similar to our own. Where our model differs is that the internal model of the economy is very limited, and that only total imports and exports are explained, not the particular level of each "ij-flow." Taplin, in his review, suggests that Neisser and Modigliani's research question could only really be answered "by linking together large, multisector models of each country." (Taplin 1967, p. 445). In an even more dismissive review their results are described as "disappointing and do not add significantly to what was already known." (Prais 1962, p. 573).

Another attempt at linking imports and exports into a global framework which was limited by the availability of satisfactory descriptions of each country's internal workings was that of Polak (1954) Working at the IMF, Polak produced a model of twenty-five major countries in a book entitled "An International Economic System." In a manner similar to that of Neisser and Modigliani, Polak has imports linked to exports in a very simple way. Income is a fixed share of exports, what he called the "export multiplier," and imports are a fixed share of income, so-called "marginal propensity to import." The latter of these assumptions is identical to how import demand is decided in our model, but the former is much improved by the input-output models have only relatively recently become available. In Taplin's review, he supports this idea, saying "Polak would have preferred to describe the transmission mechanism [linking import and exports] by linking together models for individual countries [but] such models were not — and still are not — available for a sufficient number of countries" (Taplin 1967, p. 445). Despite these shortcomings, Prais, a difficult reviewer to please, noted that "comprehensive models of this type are of interest and may provide important guiding lines for more detailed work." (p. 573).

Rudolf Rhomberg , a colleague of Polak's at the Research and Statistics Department of the IMF, and inspired by Polak to study the transmission mechanism between exports and imports in a global context, is the first author in our review whose work was taken up and developed in a sustained and co-ordinated way. As he was an IMF employee, and not an academic scholar, much of Rhomberg's early work went unpublished, but he gave a very short description in the journal *Econometrica* of the model which he and Polak had been working on since the early 1960s (Rhomberg 1966a).

This model went on to be known as Project LINK, and a full description of it is given on p. 46. For now, we will give just a brief overview of Rhomberg's 1966 review. The model takes as exogenous the level of output, prices and "certain financial variables in the United States and Western Europe" and seeks to explain, and project into the short-term future, the current account and balance of payments of these two regions and "the rest of the World." Rhomberg says that he plans to expand the model to a larger number of separate regions (he aims for eight to ten), to distinguish imports by commodity group, and to include international capital flows. He is also the first author in our study to explicitly mention development aid flows as being of interest to a model of international trade, something to which we turn our attention in other parts of this thesis.

This completes our review of what might be described as the first phase of global modelling. It includes early attempts at a multi-region input-output model for the world, simple gravity modelling, and some very simple early attempts at describing the levels of trade through descriptions of the inner workings of countries' economies. We can therefore see that, by the end of the 1960s, all the parts of a working, complex model of the global economy were in place, and the developments which followed, and which follow in this review, were theoretical and data enhancements to these ideas. The most notable features of what was to come are the introductions of nonlinearities in the relationships (see p. 56 and p. 72) and attempts to "fill in" and project forward observed quantities in the data (see p. 37).

2.5 The RAS method

Many of the authors we have reviewed so far have noted similarities between the study of global trade flows and input-output analysis, and have brought techniques from the latter to fruitful study of the former.[25] A further "import" from the discipline of input-output analysis to global trade involves the estimation of unknown entries in a flow matrix, either due to missing data in a given year, or in an effort to project the entire matrix forward into the future.

In 1960, Dr Alan Brown and Professor Sir Richard Stone started the *Cambridge Growth Project*, a multi-sector model of the British economy. Many innovations which are now central to input-output analysis, such as supply and use tables and social accounting matrices, were developed either as part of the project, or by one or other of its founders. The work of the project was ground-breaking in its level of detail and its combination of time-series and cross-sectional analyses. Indeed, the project produced a successful commercial arm, Cambridge Econometrics, which is still in operation today.[26]

Stone and Brown (1962) introduced a 'bi-proportional method for projecting matrices" which was adapted by subsequent scholars to create regional matrices from the more widely-available national ones, and to fill gaps in incomplete matrices. But its original purpose was to analyse "the intertemporal behaviour of input-output coefficients" (Parikh 1979).[27] Stone and Brown took the input-output tables from 1959 for nine European countries and extrapolated them forward to 1965. Each technical coefficient (the individual elements of an input-output matrix) is subject to two forces which lead it to change over time. The first is the "substitution effect" whereby one input is replaced with another. Parikh (1979) gives the example of oil being replaced by natural gas after the 1973 oil shock (although this is presumably a rather extreme case, and one inspired by the timing of Parikh's paper.) This involves rebalancing inputs across the rows of the input-output table. The second change to technical coefficients over time is the "fabrication effect" whereby inputs are used more efficiently as fabrication methods improve. In the original formulation of the method, this results in a simple rescaling of the columns of a flow matrix, meaning that all products are used more efficiently to the

[25] From the review above, at least Beckerman (1956), Pöyhönen (1963), Metzler (1950) and Rhomberg (1966a) independently (or at least without direct citation) noticed the similarities between global trade flows and inter-sector input-output flows.

[26] A Cambridge University document entitled "The Cambridge Growth Project 1960-1987: A Catalogue of the Collection" gives a fuller version of this back-story to the work of Stone and Brown.

[27] Parikh (1979) gives a excellent overview of the RAS method and its origins, and much of this summary is condensed from his review.

same extent as time goes on.[28] Stone and Brown then assumed that the total production of each sector (the row and column totals) were known and an iterative procedure is followed, adjusting the rows and columns of the known matrix until the required row totals are reached.

This process is called iterative proportional fitting, and works in a very simple way. First the rows of the flow matrix, X, are scaled such that each row is forced to sum to the correct row total. Then the columns are scaled with the same aim, forcing the column totals to be correct. This process is repeated, first rows, then columns, until the solution converges. At this point, we have a projected version of the flow matrix which retains some of the characteristics of the original, but with row and column totals which have been stated in advance. Technical coefficients can then be calculated in the usual way.

The method is named due to a technicality in the way Stone and Brown presented it. They denote the column vector of known row sums as u, and the row vector of known column sums as v. They then postulate as-yet unknown vectors of "fabrication effects", s, and "substitution effects," r. The two effect vectors can then be found by iteratively solving the following equations, which amount to a mathematical formalisation of the iterative proportional fitting procedure outlined above:

$$\hat{r}Xs = u \tag{2.9}$$

$$r'X\hat{s} = v' \tag{2.10}$$

where $\hat{r}$ denotes the diagonal matrix of vector r. Having solved for these effect vectors, using the known row and column totals, the updated matrix, A, of technical coefficients is related to the old, known matrix via:

$$A_1 = \hat{r}A_0\hat{s} \tag{2.11}$$

It is this formulation which gives the method its (rather obscure) name.

2.5.1 RAS applied to global trade

In the early 1960s, Jean Waelbroeck (see above) was working on a verification of Stone and Brown's RAS method in the case of the Belgian economy. He suggested to a colleague, Jean Bénard, who had been working on qualitative studies of political economy in international relations, to apply the RAS method to data on international trade.

[28] It is important here to distinguish between a flow matrix, whose elements are quantities of a sector's input or output, and a "technology matrix" whose elements are the technical coefficients. The latter is found by diving the elements of the former by the relevant row sum.

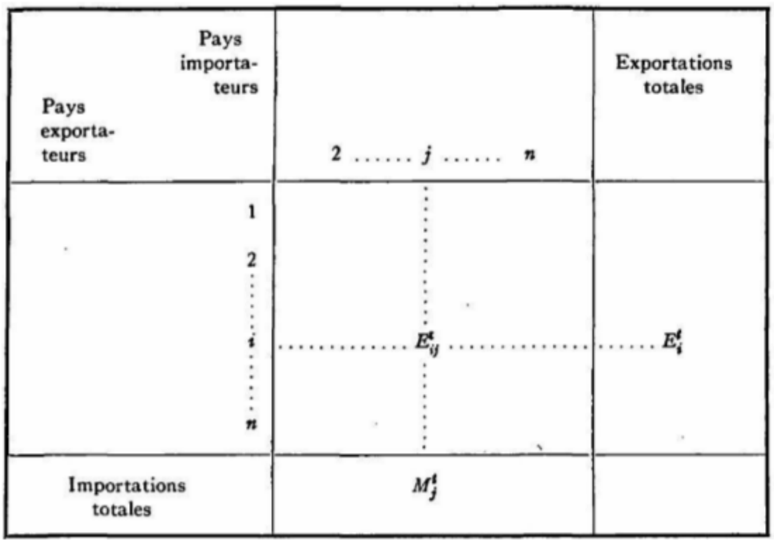


Figure 2.2: An early example of a matrix layout for performing the RAS method on the network of global trade, by Bénard (1963).

Bénard (1963) presents his trade data in matrix form, as in figure 2.2, suggesting that there are two ways to forecast future instances of the table of data.

> One might elaborate such a table with one or other of the following two methods. The more ambitious consists of predicting the volume of each trade flow according to its origin and destination by relating them back to internal economic factors in the exporter and importer...this task becomes considerable and, in the limit, begins to look a lot like the creation of a global growth model[29] (Bénard 1963, p. 542).

The second method Bénard then outlines, is that of the "*méthode de double proportionnalité*," the RAS method that we have seen above. He was, however, not impressed with the method, either theoretically or in practical terms. He admits that there is no economic justification for the procedure, but attempts a rationale by saying that "a non-negligible inertia" (p. 548) must exist whereby import and export totals must grow proportionally. He tests the squared error in each trade flow between RAS predictions and observed values from available trade data matrices and finds the results to be "very disappointing."

Despite these misgivings, he concludes that "as long as trade is at a sufficient volume, and as long as serious institutional changes do not intervene, the RAS method is a convenient and valuable procedure for projecting the trade network forward between 10 and 15 years" (Bénard 1963, p. 553). However, in their much later summary of the literature on input-output methods, Rose and Miernyk (1989) are scathing about the empirical success of the RAS method as applied to the estimation of regional tables

[29] We have translated this quotation from the original French.

from larger-scale national ones. They concur with Bénard's early findings that average errors between predicted and observed flow quantities were high enough as to be unacceptable and although the authors are keen to defend Professor Stone's considerable reputation, their dismissal of the method is expressed in no uncertain terms:

> ...the regional table bears little resemblance to a survey-based table. This should surprise no one. Stone's notions of substitution and fabrication effects are conceptually sound. The idea that these effects will operate uniformly along rows and down columns, however, cannot be supported on either theoretical or empirical grounds. Indeed, the opposite assumption that coefficient changes resulting from substitution or changes in value added will not be identical, or even necessarily in the same direction, is much easier to support (Rose and Miernyk 1989, p. 250).

Finally, Polenske (1997) gives an exhaustive list of the literature which uses the RAS technique, throughout the 70s and 80s, although her survey only features two global uses, both of which we have reviewed here. Polenske first released an article (not peer-reviewed) criticising the RAS literature in 1986. In the dedication to her 1997 topic Polenske thanks her much earlier co-authors "with whom I discussed these views in the early 1980s." That this critique only appeared in a peer-reviewed journal in 1997 tells us something about the lack of interest in this topic after the 1970s.

A lone voice advocating for the RAS technique, that of Harrigan, McGilvray, and McNicoll (1980), shows that it performs better than other non-survey methods but still concludes that there exist "significant relative differences between individual coefficients, even for the RAS technique."

2.6 Joining theoretical and empirical work

The models and methods discussed so far have been primarily ‘data-driven,’ that is, defined primarily by their relationship to observation, and not to theory attempting to explain the behaviour of economic actors. In the mid-1960s Paul Armington, a trade theorist at the IMF, wanted to bring this strand of the trade literature together with what he called the “traditional theory of buyers’ behavior” (Armington 1969a, p. 171). He describes the literature reviewed by Taplin (1967), which we have summarised here, as the “modified shares approach,” whereby it is assumed that each exporter provides a constant share of each importer’s demand but that the constant-shares matrix is adjusted to take account of various factors which either increase or decrease that share. He then asks how this approach might be justified using well-understood trade theory:

> [C]an the traditional theory of buyers’ behavior provide a satisfactory rationalization of the modified-shares approach? Can a few assumptions tie theory and practice together? This paper shows that the modified-shares approach does not require any radical departures from the traditional theory of buyers’ behavior (Armington 1969a, p. 171).

For the theory which was to join these two strands of the literature, Armington began with Hicks (1949), and what Armington describes as the “general Hicksian model.” We now take a short detour to give a brief sketch of the characteristics of the Hicksian model.

2.6.1 Hicks’ critique of Harrod

Hicks’s now widely-cited work developing a trade theory began life as a review of a model by “Mr. Harrod.” The Harrod work under review is now known as the ‘Harrod-Domar model,” since it was developed independently by the two researchers. Hicks has a great deal of admiration for the model, saying “...no one can study My. Harrod’s work at all deeply without feeling that results of really great significance are just around the corner” (Hicks 1949, p. 109) but notes that Harrod’s system is unstable mathematically: it only has one equilibrium solution for the level of output, which is zero. The alternative to this equilibrium is an explosion to infinity.

Harrod’s model centres around an investment/output equation which shows how savings and investments lead to growth in output. But Hicks shows that an increase in the rate of savings means that savings begin to exceed investment and therefore *diminishes* output. Harrod uses this mathematical quirk as a strength, saying that it explains the

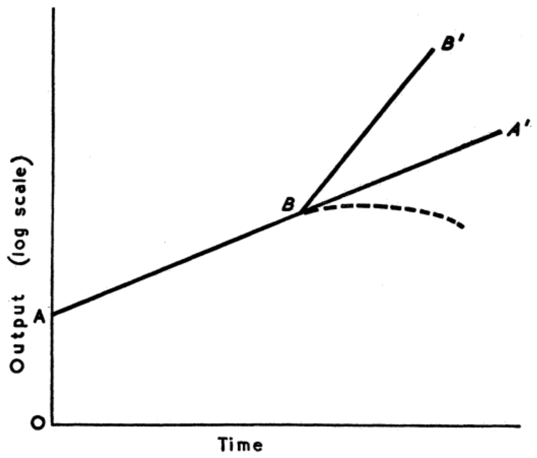


Figure 2.3: Hicks complained that Harrod's popular trade model was unstable. If an economy is travelling from A to A' and the saving rate increases instantaneously at B, the resulting path is mathematically undefined.

fluctuations seen in the real economy, but Hicks protests that "mathematical instability does not in itself elucidate fluctuation. A mathematically unstable system does not fluctuate; it just breaks down." (p. 108).

Hicks fixes this instability by introducing what he calls "frictions" into the model to prevent it going to zero or to infinity. The take the form of lagged versions of the growth variable, which turns the equations into levels into those in changes. The model is then interpreted as explaining *variations* in growth around a long-term trend. This means that the system having an equilibrium value of zero now makes perfect sense: in equilibrium there is not no growth, but rather there is growth precisely at the trend value. Hicks's improved model has fluctuations, as the real economy does, and he concludes that he has " shown that a cycle, which is strongly reminiscent of that which we experience, can be explained on the basis of a minimum number of hypotheses, each of which is very reasonable in itself" (Hicks 1949, p. 949). Almost immediately after writing the review article summarised here, Hicks began work on a book on the subject in which he mentions that a number of other authors (Leontief 1949; Goodwin 1949) were "hot on [my] heels" (Hicks 1950, p. vi). For example, the latter of these other authors writes:

> ...there is a lower limit to negative net investment given by the rate of attrition of capital stock [and an] upper limit imposed by the capacity of the investment goods trades. By combining these two considerations with a dynamical (involving a lag) multiplier, we get a single, stable, limit cycle for national income. It has an equilibrium point but this point is unstable. Therefore, the system is either always on the characteristic limit cycle, or, if disturbed continually, always approaching it (Goodwin 1949, p. 185).

2.6.2 Armington and differentiated products

Armington (1969a) introduced to trade theory a neat mathematical way of modelled the fact that products made in different countries are not perfect substitutes for one another. Thus, as he says, "French machinery, Japanese machinery, French chemicals, and Japanese chemicals might be four different products distinguished in the model." He makes an explicit, linguistic distinction between *goods*, such as machinery, and *products*, such as French machinery. We will follow this convention here.

He begins by treating every importer as identical, and that (the representative) importer's utility is a function of every good from every country (thus, in his language, of every product.) This means that demand for good i produced in country j, X_{ij}, depends on every price in the system (as well as on a budget constraint.) But this makes the whole system hard to handle using the computational power available at the time.

He instead finds a way to "collapse" this dependence on every price by separating the total demand for a *good*, from the make-up of *products* which satisfy that demand. The importer's utility depends only on the total amount of each good imported. Thus, in utility terms, the importer does not care where these goods come from. But the shares of each country of origin in this total import is determined separately, and in a way that is fixed per good (see eq. 4, p. 164).

Using a result from Solow (1955), Armington shows that this is equivalent to assuming that "the marginal rates of substitution between any two products of the *same* kind must be independent of the quantities of the products of all *other* kinds" (Armington 1969a, p. 164). In other words, the extent to which you are willing to swap your Italian leather for Japanese leather is unrelated to your relative consumption of all other, non-leather, products. Decisions in one product market are not affected by purchases in other markets. Armington calls this "the assumption of independence."

Solving the model is now a two-stage process: First find the total demand for each good. Then divide that demand up among competing exporters. In the latter stage, market shares are assumed to depend only on relative product prices (not on the total size of the market.) Futhermore, within the market for a given good, each product has the same elasticity of substitution with each other product. These elasticities are constant and do not depend on the size of the market.

The total export demand for a given good on any exporting country (recall that all countries are the same) is therefore a constant elasticity of substitution (CES) function of demands for each of that country's products. Armington expresses this as:

$$X_i = \left[\sum_{j=1}^{m} b_{ij} X_{ij}^{\rho_i} \right]^{-\frac{1}{\rho_i}} \tag{2.12}$$

with X_i being total trade in good i, and X_{ij} being the demand for good i as produced by country j. The model is completed with a description of how relative prices affect demand for products as follows:

$$X_{ij} = b_{ij}^{\sigma_i} X_i \left(\frac{P_i j}{P_i} \right)^{-\sigma_i} \tag{2.13}$$

where P_i is the average price level across all exporters and P_{ij} is the price of good i from country j. It then remains only to find $2m$ elasticities, m being the number of goods: the elasticities of demand and of substitution.

Armington finally assumes that domestic goods compete exactly as foreign goods on the global market. This last assumption is relaxed in section 6 of an empirical paper, Armington (1969b), published in the same year. In this, he applies his model to estimating the effect on the trade network of a single exogenous price change of one product. He assumes an elasticity of substitution between different countries' goods, and an elasticity of demand overall.[30] Using trade shares (from data) and elasticities (assumed) he is able to simulate a change in price and assess the effects on all the levels of trade. Armington makes the point that the share of trade observed in the data captures all the unknowable subtleties of the relationships between the various countries, such as trade restrictions, long-term contracts and traditional loyalties. This is very much in the spirit of a lot of the modelling philosophy outlined in this thesis.

These two works (Armington 1969a; Armington 1969b) were hugely influential in their ability to trace the effects of price changes though the whole trade network, using the whole matrix of trade flow data, rather than looking at countries pairwise in isolation. One or other of the papers is cited in a great deal of the general equilibrium work which followed. See, for example, Miller and Spencer (1977) for a relatively early general equilibrium (GE) approach, Dixon and Parmenter (1996) for a computable GE approach, Hertel (1997) for an overview of the GTAP project (of which more below).

[30] When a price increases, buyers will switch to other sellers. Given that the buyer did not choose these sellers for this trade before the price increase, it follows that the new scenario will end up less favourable to the buyer than the previous one. Therefore the buyer should be expected to purchase less of the good overall. The switching and the changing of volume bought are known as the substitution and wealth effects respectively. In Armington's paper, the substitution elasticity is assumed to have a value of 3, and the price elasticity a value of 1. These are "roughly the orders of magnitude that might be expected on the basis of past research and a priori reasoning" (p. 182).

Feenstra (2015) mentions Armington, but only in the context of his distinction between goods and products (see above). But the first users of Armington's formulation were a group attempting to build a fully-fledged model of the global economy with internal economic descriptions linked via a model of the trade network. The first published outputs of this project, Project LINK, were gathered into a book edited by Ball (1973), to which we now turn.

2.7 Project LINK

In 1968, a group of researchers at the Committee on Economic Stability, based in the Social Science Research Council in the USA, proposed a project to study what they called "international transition mechanisms," in other words, mechanisms by which domestic economic fluctuations could result in changes to international trade flows. The idea was that national models, which had always had an internal economic description and some simple trade equations, could be linked up to one another by harmonising the trade equation part of each model. The project went on to run for many years and was called Project LINK.

The first published description of the aims of this ambitious multinational project were given in 1970 by an IMF economist, Rudolf Rhomberg, whom we mentioned earlier in this review, who proposes the "construction of a model of trade and financial flows as a centerpiece of the complete world model and the linkage of each national model to this central trade and financial model" (Rhomberg 1970, p. 2). Although this is the first mention of the project by name, Taplin (1967) attests to the fact that work on world trade models had been going on at the IMF since at least 1962 and project LINK seems to have this work at its heart.

Shortly after Rhomberg's presentation of project LINK, an entire book was released dedicated to the subject, entitled *The international linkage of national economic models*. In his overview of the book, editor Robert J. Ball described the aims of the project as being to "explain" the level of imports using national models, disaggregated "according to certain SITC groupings," which would then be used to find "consistent world solutions [which] preserved accounting identities between exports and imports" (Ball 1973, p. 1). He goes on to give a statement of the democratic, decentralised *raison d'être* of the project:

> A guiding philosophy in LINK has been that each model builder knows his own country best. The project was not based on the standardised construction of national models and trade equations at one remote research centre, but on the effective integration of large existing models with some international standardisations acceptable to all participants (Ball 1973, p. 2).

A number of developed countries (the USA, Japan, UK, Italy etc.) were included, and socialist and developing country models were provided by analysts at the United Nations Conference on Trade and Development (UNCTAD), by the IMF, and by Stanford University.

Many of the chapters in this book went on to be very widely cited in their own right. We will look in detail at some of the most important of these.

2.7.1 Rhomberg's appeal to Armington

The first model in the book is that of Rhomberg (1973), which seeks (as do most such models) to determine the quantities and prices of internationally trade goods. He begins by critiquing the classical, Walrasian, approach of using supply and demand schedules which begins by determining whether countries demand more or less of the goods that they produce. This approach, he claims, should lead to countries being either importers or exporters of each good, but not both, and this is not what we observe in the data.

Furthermore, a Walrasian approach would not allow for the differentiated prices around the world seen in the data, beyond via transportation costs and other trade barriers. It would nevertheless require the model to have a demand and supply equation for every price in the world which would, at the time, have been unworkable: "[c]learly some fairly drastic simplifications must be introduced to make this system of equations manageable for either econometric estimation or policy simulation" (Rhomberg 1973, p. 15).

Rhomberg appeals to Armington's assumptions in the following terms: Italy's demand for machinery can depend only on Italy's income and on "world price levels" (somehow defined) for machinery. And Italy's demand for machinery from the UK depends on the ratio of the UK price to all other prices, including Italian prices. With Armington's assumptions of constant elasticities of substitution between products of one commodity class, which are the same between all pairs in the class, Rhomberg notes that demand for each product is then a constant fraction of that country's demand for the good in question "modified by the price term just mentioned [the ratio of local price to world price]." In short, Rhomberg proposes linking the country models available to Project LINK with an Armington model of global trade.

2.7.2 Hickman and Lau

In his contribution to the book on Project LINK, Hickman (1973) outlines a modelling approach which was then operationalised for empirical estimation in Hickman and Lau (1973) which models world trade "using the trade shares matrix approach." This latter paper went on to be hugely influential in the trade literature. Indeed, Alexander Italianer, in describing the CES literature as part of his much broader review of trade literature, chooses Hickman and Lau (1973) as the definitive statement of CES modelling, saying "the others may be considered to be closely related to, or copied from, their paper" (Italianer 1986, p. 69).

Hickman and Lau (HL) use much of the theoretical work and assumptions of Armington (1969b) but they use it to develop linear equations which enable them to estimate elasticities of substitution between different exporters' products, rather than simply selecting values for them as Armington does. They can then specify demand functions for each of 25 industrial countries (along with a socialist bloc and "the rest of the world") which can be used to forecast trade values beyond their sample. They start by assuming that total imports have already been determined and focus on the allocation decision between competing exporters. Will we reproduce the key parts of their methodology here. Note that where equations are reproduced from Hickman and Lau, we have retained the original equation numbers and prefaced them with *HL*.

HL first present a first-order linearisation of Armington's CES specification which we reproduce here for convenience. For each importer, define an import index for country j (m_j^*, proportional to the actual quantities purchased, M_j) with constant elasticity of substitution σ_j. Thus, for importer j, any two pairs of exporting countries have elasticities of substitution σ_j. For notational convenience, introduce $\rho_j \equiv 1/\sigma_j - 1$ and the CES function can be expressed compactly as:

$$m_j^* = \left[\sum_{i=1}^{n} a_{ij} x_{ij}^{-\rho_j} \right]^{-1/p_j} \tag{HL3}$$

where x_{ij} is the import arriving at country j from exporting country i. They begin by rearranging this expression, along with some useful algebraic simplifications, to obtain an expression for the (cost minimising) import quantities which, with a sum over goods categories in the denominator, has a tricky nonlinear form:

$$x_{ij} = \alpha_{ij}^0 \left(p_{ij}^x\right)^{-\sigma_j} \left[\sum_{k=1}^{n} \alpha_{kj} \left(p_{kj}^x\right)^{-\sigma_j} \right]^{-1} m_j \tag{HL15}$$

where a zero in the superscript denotes a value from a base year in which all export prices are equal to unity.[31] α_{ij}^0 is the base year share of imports from exporting country i, x_{ij}^0/m_j^0, and p^x are export prices.[32]

[31] Since the authors are only interested in elasticities, i.e. responses to price *changes*, they find it convenient to imagine a base year in which all prices are unity. Scenarios are then be developed in relation to this base year.

[32] Note that we have made a slight change to the notation of Hickman and Lau (1973). The use of p^x and p^m as export and import prices respectively leads to an unfortunate notational clash when these

This is then linearised twice using first-order Taylor approximations. The first is around $p^x_{ij} = 1$ for all i and j, and the second is around $m_j = m^0_j$. This leads to a much simplified expression for the export quantities, which is linear and estimable from easily available trade data on quantities (and hence shares) and prices:

$$x_{ij} = \alpha^0_{ij} m_j - \sigma_j x^0_{ij}(p^x_{ij} - p^m_j) \quad \text{(HL21)}$$

where p^m_j is an index of import prices, weighted by the share of each exporter in total imports.

In a final step, HL allow for a constant time trend in the original CES formulation, which reflects "taste changes over time" (p. 353). A similar linearisation procedure is then followed resulting in the final equation to be estimated:

$$x_{ij} = \alpha^0_{ij} m_j - \sigma_j x^0_{ij}(p^x_{ij} - p^m_j) + \sigma_j x^0_{ij} r_{ij} t \quad \text{(HL28)}$$

where t is a time trend and each r_{ij} is a constant.[33]

Matters are then somewhat complicated by the authors with the introduction of the idea that actual trade responds slowly to changes in prices. They call this graduated response "adjustment lags." They find that introducing a lagged adjustment is theoretically inconsistent with the criterion that trade shares must sum to unity (what they refer to as "the adding up criterion.") Instead they opt for a slowly-changing expectation of the price level.[34] This theoretical addition is not rigorously justified other than by saying that "[i]mmediate revision of price expectations seems implausible to us as an assumption in behavior, so we prefer the dynamic model on intuitive grounds." (Hickman and Lau 1973, p. 368). But it was presumably initially included to improve the goodness-of-fit of the final regressions: adding an additional degree of freedom means that trade shares are not expected to respond 'perfectly' to changes in price, but rather to move slowly in one direction or the other in response. This gives HL the ability to use a time

quantities are raised to a power. Therefore, wherever we need to express one of these raised to a power, we have added parentheses to clearly separate the exponent from the non-mathematical x or m notation.

33 The r_{ij} are actually algebraically related to the import shares and are therefore constrained (since shares must sum to unity.) This additional constraint on the estimation is mentioned on p. 362 of Hickman and Lau (1973).

34 The adjustment lags discussed in this section are all of the form $z_j = \delta(z_j)^* + (1-\delta)(z_j)_{-1}$, $0 < \delta \leq 1$, where z_j is the per-importer variable of interest and z^* is the desired or expected value of that variable in the current period. Thus the realised values of z are 'smoothed' by the previous period's z and an expected or desired current value. Equations (HL36) and (HL39) are examples of this.

series of trade data to estimate the elasticities of substitution.[35] They do this using an IMF series, provided by Taplin, for the years 1961 to 1969 for 27 countries or regions.

They then present a table of the estimated substitution elasticities for each country, finding that estimated elasticities range between close to zero and 15, with an average of 3.03. They are clearly pleased with this result, comparing it to a section in Armington (1969b), noting that "a value of 3 was assumed for all markets by Armington in an empirical application of his original model, on the grounds that the figure was roughly of the order of magnitude to expected from past research and a priori reasoning." (Hickman and Lau 1973, p. 367).

They conclude by attempting to predict trade flows from the empirical elasticities (and price adjustment parameters) finding, unsurprisingly, that the model does far better in-sample than out-of-sample. Nevertheless, the out-of-sample estimates are generally impressively close to data values, in part thanks to the dynamic (slow-changing) nature of the model. As long as predictions are not made too far ahead, and nothing too dramatic happens between the end of the sample and the start of the prediction, this kind of dynamic model can be expected to produce reasonably good predictions or, at least, predictions which are reasonable.

An early application of the work of Hickman's work for Project LINK is an attempt by Marwah and Diwan (1975) to separate market and non-market effects when looking at trade with a particular focus on developing countries/regions. They are interested in quantifying "implicit subsidies or grants in international trade." By this they are referring to the fact that countries which export a lot get an advantage in trade which goes beyond a mere price signal. Indeed, although the authors do not mention it, the gravity model discussed above makes this explicit: larger countries export more, due to either economies of scale or, perhaps, a greater market-making power. In general, trade between "unequal partners" should be expected to face market distortions in favour of the dominant partner. Even development aid is treated by Marwah and Diwan as an implicit subsidy to the donor country since, at the time at least, aid flows were closely tied with improved trade terms for the donor country and hence constitute an "implicit subsidy" to the donor country.

They develop a version of the Armington/Project LINK CES formulation with an additional non-market gravity-style term accompanying a price-ratio term as follows:

$$a_{ij} = a_{ij}^{0}(\text{price ratio term})^{\alpha_{ij}} \left(\frac{X_i}{M_i}\right)^{\beta_{ij}} \times \text{drift term} \tag{2.14}$$

[35] They also estimate the adjustment parameters which allow actual trade quantities to adjust slowly in response to changes in price.

By appealing to a RAS methodology, they calculate the trade which might be expected under proportional trade shares. They then split the resulting differences from proportional trade into the market effect, governed by price ratios, whose impact is moderated by the α parameters, and all other (hence non-market) effects, whose impacts are moderated by the β parameters. Allowing for a drift parameter to account for changes in, e.g. taste, over time, they are able to estimate the sizes of β and hence, per pair of trading countries, the extent to which non-market forces are boosting or suppressing trade.

By focusing on trade to and from North America, they find that Western Europe contributes the majority of the implicit subsidies given to North America and most of North American implicit subsidies are given to Latin America and Japan. Although they don't state this explicitly, presumably the authors are hoping to argue that North America's trade favours Latin America and Japan over other developing regions such as the rest of Asia and Africa. Although this work was very little cited, Marwah's work on developing countries seems to have inspired a much more widely-read work on bilateral trade elasticities by Marquez (1990) which we look at on p. 71.

2.7.3 Klein and van Peeterssen

In their contribution to Ball (1973), Klein and Van Peeterssen (1973) seek to operationalise the theoretical model-linking work elsewhere in the book, accounting for problems of lack of data, imprecisely recorded data, small samples and "inherent collinearity" which stand in the way of obtaining reliable estimates of the parameter values which make up a particular model specification.

They present three specifications by which to link the various national and bloc-level models in the project. The specifications are named "Mini-LINK," "Midi-LINK," and "Maxi-LINK" and the authors briefly outline each in turn. (We maintain the equation numbering of Klein and van Peeterssen (hereafter 'KP'), preceding each equation with 'KP'.)

The authors describe how, in a perfectly specified global model, every flow X_{ij}^{k} of commodity k from country i to country j would be explained and global imports would sum to global exports. This would hold both for each separate good and, by extension, in aggregate. The Mini-LINK is restricted to ensuring that total imports M_i sum to total world trade. Each LINK model has, by design stipulation, an exogenous variable for the level of world trade, WT. The idea behind Mini-LINK is to find a level for WT, total world trade, which, when fed into the country models as an exogenous variable,

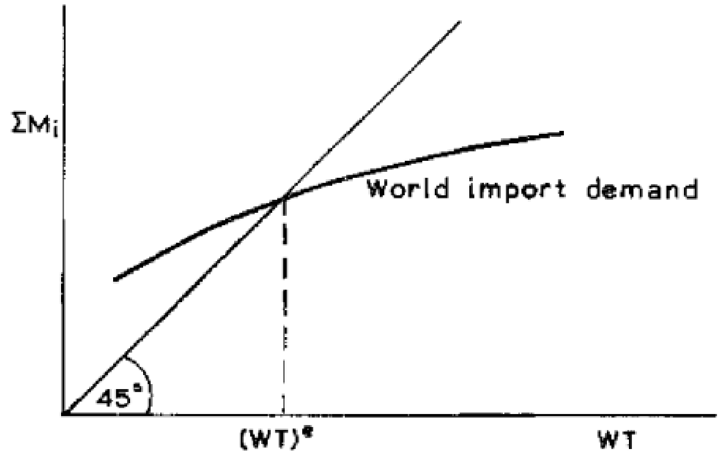


Figure 2.4: Klein and van Peeterssen's graphical method for solving the "Mini-LINK" such that total modelled import demand equals world trade.

produces import demands such that:

$$\sum_i M_i = WT \tag{KP2}$$

This is done first graphically, as shown in figure 2.4. KP "request each model builder to solve [their model] 4 times, each with a different assumed level of world trade" (Klein and Van Peeterssen 1973, p. 435). The resulting import demands are then summed across all the models. These four total import demands are then graphed against the values of WT that generated them, and the 45° intersection is taken as the equilibrium value.

But KP also outline a more dynamic way of solving for the equilibrium level of world trade. They outline an iterative procedure which first sets world trade and then calculates the models. If total import demand exceeds world trade, then the models are solved again with world trade assuming the new value for total import demand. This is repeated until equilibrium is reached.[36]

KP immediately criticise this method for not allowing a price-adjustment mechanism to play a role in reaching a global trade equilibrium, but nevertheless conclude that "this simple search for a balance in [KP2] is a powerful equilibrating tool and should lead us to a better solution fo the world economy that gives us a better estimate of world

[36] This system is quite similar to the iterative solution scheme used throughout this thesis, outlined in chapter 3. We discuss in that chapter why this simplest solution scheme is sufficient in our case to demonstrate the kinds of global modelling techniques we are proposing. In fact, our model is a combination of the Mini-LINK with its simple iterative solution method, and the Maxi-LINK, with its focus on making global exports equal global imports.

trade than the summation of uncoordinated solutions across countries." (Klein and Van Peeterssen 1973, p. 439).

The next solution scheme to be outlined, the "Midi-LINK," adds price adjustments to the equilibrating scheme. A difficulty is immediately encountered, in that "[a]ll country models are not equally complete in the use of price variables...but there are enough price variables in the LINK system to make a price adjustment process worthwhile." The Midi-LINK adds an iterative process to solve for import prices, which are exogenous in the country models. This endogeneity presents an accounting problem, because export prices are endogenously defined in the country models. This means that as the country models are solved, import prices (exogenous) will tend to drift away from export prices (endogenous), although it is a simple accounting identity that these are one and the same set of prices, since one country's export is another country's import.[37] KP introduce the simple assumption that (until now exogenous) import prices will move at the same rate as an index of export prices, weighted by total export volume, similar to that used by Hickman and Lau (1973). The new system then allows forecasts of both trade volumes and changing global prices. How a country responds to changing prices is, of course, up to the country model. In this way, the global part of the model can be very simple, while allowing for non-linearities and complexities within the country models.

The final version of KP's model-linking system, the "Maxi-LINK" adds the further accounting constraint that global exports must equal global imports. They move from an import-constrained system:

$$\sum_i M_i = WT$$

to one where the full accounting identity of global trade is respected:

$$\sum_i M_i = \sum_i E_i = WT$$

This seems like an obvious move, but because exports are determined endogenously by the country models, it is not a given that they will sum to the same total value as imports, which are determined using either the Mini- or Midi-LINK such that imports sum to world trade. The Maxi-LINK determines imports from the country models, according to behavioural equations, then determines export levels such that exports match imports. This is done in such a way as to preserve, as far as possible, a matrix

[37] As KP note, this equality of prices holds up to a set of "c.i.f.-f.o.b. differentials" whereby import and export prices differ by carriage-related costs borne either by the exporter (c.i.f.) or the importer (f.o.b.).

A of trade shares. In fact, the reality is slightly more subtle, and A is allowed to "drift" over time from its empirical values.

Project LINK remained active throughout the 70s, with much work being done on forecasting the trade share matrix drift outlined above. Hickman and Schleicher (1978) and Beaumont, Prucha, and Filatov (1979) are early examples of this. Both of these use a method called the "Linear Expenditure System," due to Klein and Rubin (1947) and Stone (1954), to do the trade share forecasting.[38] We will next give a brief outline of this system, as it went on to be very widely used in the global modelling literature.

2.7.4 The Linear Expenditure System

The Linear Expenditure System (LES) was designed to express a cost-of-living index "in terms of measurable phenomena which are independent of the subjective concepts of utility" (Klein and Rubin 1947, p. 84). In simple terms, utility is eliminated from demand equations by solving them assuming that utility is maximised and, at the lower-bound of just being able to afford a consumption bundle, constant. By first assuming that utility has been maximised for, and then holding it constant, we can express demand in some future period as a function only of prices and income in an observed base period and prices and income in the forecast period. Stone (1954) completes the picture by taking this theoretical work and showing how these forecast demand equations can actually be estimated.

In the exposition of Pollak and Wales (1969), a specific utility function is assumed for concreteness, with a simple log form:

$$U(X) = \sum_{k=1}^{n} a_k \log(x_k - b_k)$$

The a_k are weights for how important good k is for the consumer's utility, and the b_k give "necessary quantities" for good k, about which the consumer has no choice. With the usual budget constraint, μ, and maximisation they derive the demand functions, and the resulting expenditure function for the ith good looks like:

$$e^i(P,\mu) = p_i b_i + a_i \left[\mu - \sum_{k=1}^{n} p_k b_k \right]$$

[38] The Linear Expenditure System was invented by Klein and Rubin (1947), the same Klein as we are discussing here, and first estimated by Stone (1954). These are both dense mathematical treatments, but Pollak and Wales (1969) give a much clearer overview of the method. Stone (1954) is the most widely-cited of the three papers, by some margin.

Thus expenditure is split into a required, inelastic part, $p_i b_i$, and a "supernumerary," or discretionary, part, given by $\mu - \sum p_k b_k$, the remainder available from the budget μ after the necessary expenditures.

The remainder of Pollack and Wales's outlining of the linear expenditure system consists of two simple specifications for how b might change over time, first linearly:

$$b_{it} = b_i^* + \beta_i t$$

then via a habit-forming mechanism whereby future minimum consumption levels of good k depend on the current level of consumption of good k:

$$b_{it} = b_i^* + \beta_i x_{it-1}$$

However the dynamics of b are modelled, estimating the linear expenditure system is as simple as finding values for the a and b for each product. This can be done using a time series of prices and consumption levels, which Pollak and Wales (1969) do for four categories of good–food, clothing, "shelter" and "misc"–for the years 1948 to 1965.

This concludes our look at a number of the works associated with Project LINK and at the linear expenditure system. As we shall see, many aspects of this work went on to be included in large scale models which used data for estimation and attempted to estimate both trade flow magnitudes *and* the changes in price associated with changes in those flows. In this latter sense, such models are known as general equilibrium and it is to these that we now turn.

2.8 Johansen, CGE and GTAP

The linear expenditure system outlined above, for estimating the expenditure in models of the macroeconomy, became standard in models which combined neoclassical economic models of demand and price with data-driven methods such as input-output and the use of trade shares matrices. This latter modelling tradition went on to be known as Computable General Equilibrium (CGE), and began in 1960 with a model of the Norwegian economy made by Leif Johansen. Before describing Johansen's model and CGE in general, we will take a short detour through social accounting matrices, a way to organise thinking about intra-economy flows due to Stone and Brown (1962).

2.8.1 Social accounting matrices

Solutions to matrix equations via matrix inversion would go on to be crucial to Johansen's approach to solving multi-equation systems simultaneously. To see how Johansen organised his equations into a matrix, it is useful to follow Robinson (2006) in beginning with an outline of social accounting matrices (SAMs) upon which, says Robinson, CGE models are "always based".[39] Robinson presents a neat diagrammatic formulation of the SAM, reproduced here in figure 2.5. This is accompanied by several definitional accounting identities, which amount to the same as saying that each row must add to the same as its accompanying column:

$$
\begin{aligned}
GDP = X + T^X &= D + E \\
D + M &= C + G + I \\
Y &= X \\
Y &= C + T^H + S^H \\
T^X + T^H &= G + S^G \\
I &= S^H + S^G + S^F \\
M &= E + S^F
\end{aligned}
$$

These equations can also be characterised qualitatively by looking at each of the "accounts" (row/column headers) in the matrix as being an "agent." This agent can be though of as either buying (columns) and selling (rows) "things", or as giving out (rows) or receiving (columns) money. These two flows, things and money, are inverses of one

[39] Note that large sections of Robinson (2006), including the diagram we reproduce here, are lifted directly from Robinson (1989) and the choice to cite one work rather than the other makes little difference to the exposition we present here.

	Activities	Commodity	Factors	Hshld	Govt	S-I	World
Activities		D					E
Commodity				C	G	I	
Factors	X						
Household			Y				
Government	T^X			T^H			
S-I				S^H	S^G		S^F
World		M					

Definitions:

D: production sold domestically	C: consumption
E: exports	G: government demand
X: production (GDP at factor cost)	I: investment demand
T^X: indirect taxes	S^H: household savings
T^H: direct taxes on households	S^G: government savings
M: imports	S^F: foreign savings
Y: factor payments to households	S-I: savings-investment account

Figure 2.5: A diagrammatic representation of a simple social accounting matrix (SAM) taken from Table 11-1 of Robinson (2006). The matrix is described by Robinson as "macro" because it excludes the intermediate inputs matrix (input–output table).

another. Thus, each entry in the SAM is both a flow of money from column to row and a flow of things (very loosely defined) from row to column.

For example, looking down the first column, Factors receive money X from Activities. In exchange, those Activities receive productive output (a "thing") from Factors, of value X. Similarly, the Government receives money T^X in taxes from Activities. In exchange, Activities receive T^X-worth of government services, such as infrastructure, a peaceful, educated society etc.[40] Then, looking at the first row, Activities distribute this productive output and these government services among the Commodities account (i.e. domestic production) and the World account (i.e. exports.) Thus, we arrive at the first of our accounting identities, $X + T^X = D + E$.

As a second example, the Commodities account transfers money, in the first column, to economic Activites (D) and to the World (M) and receives consumable commodities in exchange. These commodities are then distributed among Households, C, the Government, G, and the savings-investment account, I. In turn, money is exchanged by each of these latter three accounts into the commodity account. That these money flows must balance the money transferred by the Commodities account to the Activities and World accounts is our second identity. The SAM can thus be seen not as a model

[40] Precisely what is received by economic activities from the government in exchange for its taxes is clearly a complex and controversial topic. We do not investigate this topic further other than to say that "something" is provided to Activities by the Government in exchange for taxes.

per se, but as a description of flows of money and things around the economy between various actors, some of which are real and others of which are theoretical. How the actual balances are achieved (for example, how precisely the money flow to Factors is converted into Commodities) is the responsibility of a model, which may use the SAM as a framework for thinking about which flows and actors need "explaining." As with input–output modelling, where final demand is exogenous and all other entries in the matrix are determined endogenously, a SAM model must select which of its variables are to be exogenous. This choice, between exogenous and endogenous accounts, detmerines the "closure" of the model (Robinson 2006, p. 209).

The simplest possible model with a SAM framework is a Leontief-style one in which actor behaves according to fixed column coefficients, exactly as is done in input–output modelling. Thus, Commodities use domestic production, D, and imported production, M, in fixed proportions. And Households have a fixed propensities to consume, C, to save, S^H, and to pay taxes, T^H, as a proportion of their total income. This proportional response to an exogenous change, or shock, means that the change in the level of the economy overall is generally greater than the initial shock. For this reason, proportional models of this nature are called SAM-multiplier models. It is also for this reason that these models are often described as Keynesian: output is determined by demand and it can be shown that the multiplier on the response is the same as the Keynesian demand multiplier (Robinson 2006, p. 210).

Notice that, until now, no mention has been made of price flexibility, supply constraints or time dynamics. By adding a description of the production process, and a mechanism by which price can respond to changes in, for example, demand, we arrive at a version of Johansen's model, often cited as the first computable general equilibrium model. But Johansen himself did not use that term. We will therefore outline Johansen's model in sketch form, and then look in more detail at CGE modelling more generally.

2.8.2 Johansen's Norwegian Model

Leif Johansen was an economics student at the University of Oslo in the early 1950s, and worked with many of the names we have covered here in this review, most notably Richard Stone, whose linear expenditure system we saw above, and Jan Tinbergen, on whose work much of the subsequent global economic modelling literature is based.[41]

Johansen studied the relationship between industries, for the first time expanding growth analysis to situations of more than two sectors. He was particularly interested

[41] A good overview of Leif Johansen's early life and the context in which he did his groundbreaking work is given by Bjerkholt (2009).

in "deviations from uniformity in the growth process" (Bjerkholt 2009, p. 112) between different sectors of the economy. Until that point, growth had been studied as a single process in the economy, the implicit assumption being that all sectors grew proportionally with one another. Johansen's interest in Leontief's multi-sector work led him to want to build a multi-sector version of his contemporaries' growth models and two-sector models.

Despite being the originator of a huge strand of the economic modelling literature, Johansen was always happy to be self-critical and to acknowledge the limitations of his work. This quote from the late 1960s is typical of his style:

> It must be admitted already at the outset that we are aware of many deficiencies of the model, some of which could be remedied by more work and/or by better statistical data. The team which has been working on the project is quite small. There are however also some weak points concerning which we would not know exactly how to proceed in order to improve the model even with more work and better data. (Johansen, Alstadheim, and Langsether 1968, p. 71)

In the overview of his work which follows, we draw on summaries by Johansen himself (Johansen, Alstadheim, and Langsether 1968), Dixon and Parmenter (1996), Bjerkholt (2009) and Dixon and Rimmer (2016).

The model has 20 sectors, each producing according to a Cobb-Douglas production function with labour, N, capital, K, and technological progress, A:

$$X_i = A_i N_i^{\gamma_i} K_i^{\beta_i} e^{\varepsilon_i t}$$

The intermediate inputs required to produce X_i were determined by technical coefficients precisely as in input-output modelling:

$$X_{ij} = a_{ij} X_j$$

which, in some versions of the model, were allowed to change over time using 'the RAS-hypothesis of the well known Cambridge project "A Programme for Growth"' (Johansen, Alstadheim, and Langsether 1968, p. 79). and imports were a fixed (per-sector) proportion of total production:[42]

$$M_i = \mu_i X_i$$

[42] Johansen, Alstadheim, and Langsether (1968) does not give any special name to these import ratios, calling them merely "constant coefficients."

Capital, which came in two forms, broadly buildings vs. machinery, depreciated and grew with investment exactly as it would in a Solow-type growth model.[43] This capital had to be maintained against depreciation proportional to existing stocks, and capital stocks increased exogenously. Prices were determined via profit maximisation and the overall wage rate was the numeraire. Wage rates in the different sectors differed from the overall rate by fixed proportions. Production goods could go to one of four destinations: intermediate input, consumer demand, government demand and exports. Private consumption was modelled via a representative consumer maximising utility in the usual neoclassical way (an increasing utility function with decreasing returns). The size of the labour force was exogenous, and full employment was assumed. Johansen himself summarised his model neatly, as follows:

> The model can be characterized as being a disaggregated neoclassical growth model. It is constructed on the basis of an input-output description of the economy, but it allows for substitution between labour and capital; furthermore, it treats price trends as endogenous in the model, and accordingly allows for effects of changes in prices upon the direction of development of private consumption demand. Implicitly the model assumes full employment...exports and competitive imports are assumed to be exogenously determined (Johansen, Alstadheim, and Langsether 1968, p. 72).

Johansen's system had 86 equations and 132 variables. As was mentioned above, all 86 equations were arranged into matrix form, and solved using a single (86×86) matrix inversion. As Dixon and Rimmer (2016) (DR) point out, the solution of the system was a matrix containing the sensitivity of every endogenous variable with respect to every exogenous variable. They point out that Johansen faced a huge problem in interpreting the 3956 entries in this matrix, saying that he was united with current CGE modellers in finding that "it is easy to become overwhelmed and confused when confronted with a huge number of results" (Dixon and Rimmer 2016, p. 425). Johansen dedicated a good amount of his book to coming up with strategies for coping with the huge number of results his model produced. DR say that, although the exact specifications of the 1960 model is "of only historical importance," the model is still highly relevant to modern CGE modelling in two ways:

> The first is the way in which Johansen presented his model as a rectangular system of linear equations in change and percentage-change variables and solved it by matrix inversion. The second, and related aspect, is how he used the linear

[43] The two types of capital were assumed to remain in fixed (per-sector) proportion with one another. This proportion is estimated by Johansen, Alstadheim, and Langsether (1968) with results for each sector given on page 92.

representation and the linear solution method: to clarify properties of the model; to elucidate real world issues; and to check the validity of the model (Dixon and Rimmer 2016, p. 424).

2.8.3 Computable General Equilibrium modelling

With production, intermediate and final consumption, investment, prices, imports and exports all being determined within Johansen's model, and with all these things being based on empirical observation of a particular, real-world economy, this model is recognised as being the first computable general equilibrium (CGE) model.[44] Dixon and Parmenter characterised CGE modelling as being data-driven and empirical, distinct from other, more theoretical branches of the economics modelling literature: "[I]n CGE modelling the database and numerical results are intended to be more than merely illustrative. CGE models use data for actual countries or regions and produce numerical results relating to specific real-world situations" (Dixon and Parmenter 1996, p. 6).

As we saw in the sections above, social accounting matrices become CGE models once production and prices, the supply side, are added. In its simplest form, this can be done by adding a production function, usually log or Cobb-Douglas, and by calculating prices relative to a numeraire. This makes simple CGE models inherently neutral with respect to the role of money: the actual price level must be set exogenously.[45] In these cases, markets are assumed to clear, and all factors are fully employed. Extensions to this simple model are what make up the majority of the subsequent CGE literature, and are given a good overview by Robinson (2006). Beyond extending the model, the other strand of the CGE literature consists of finding alternative methods of solution to the matrix inversion approach of Johansen.

DR (Dixon and Rimmer 2016) give a short overview of the strands of the CGE literature which do *not* have their roots in Johansen's work, mentioning Herbert Scarf and Dale Jorgenson among others. But according to DR, Scarf's work developing a solution algorithm for CGE models "was largely abandoned in the 1980s in favour of much simpler methods" (Dixon and Rimmer 2016, p. 428).[46] They also imply that Jorgenson's

[44] As Bjerkholt (2009) points out, CGE is also known in some traditions as applied general equilibrium, AGE. He also mentions that the models were called "Johansen-style models" in Australia until at least the early 1980s. In parts of the outline of the GTAP model below, we will use the acronym AGE where it appears in quotations.

[45] One other, more technical, way to describe this neutrality towards the role of money is to say that the supply and demand equations are homogeneous of degree zero with respect to prices. In other words $f(kp) = f(p)$ for any constant k, for both production and consumption equations.

[46] Dixon had long been scathing about Scarf's solution method. A version of this critique appeared two decades early by the same author: "Johansen had already solved a relatively large CGE model by a simple, computationally efficient method well before the Scarf algorithm was invented. Scarf's tech-

work was confined largely to future work involving Jorgenson himself as a co-author. Only the World Bank work of Adelman and Robinson (1978) and Johansen (1960) are said to have truly influenced the CGE modelling community, and each has its own special-purpose software for finding solutions, both of which are in widespread use today. For solving Johansen-style models involving matrix inversion, the software is called GEMPACK (Pearson 1988). And for the solution of Adelman–Robinson World Bank models, the software is called GAMS (Condon, Dahl, and Devarajan 1987).

2.8.4 Adelman–Robinson: CGE at the World Bank

In contrast to Johansen's neat, computationally efficient method of solving models with a single, huge matrix inversion, the World Bank modelling work of Adelman and Robinson take a more pragmatic, flexible, but computationally more expensive, approach which "involves treating a CGE model as just a collection of non-linear algebraic equations and attacking them directly with numerical solution techniques" (Robinson 1989, p. 889)

Our brief description here uses both the original paper, Adelman and Robinson (1978), and a later summary by three World Bank economists including Sherman Robinson himself (Dervis, de Melo, and Robinson 1989), hereafter DMR.

The first way in which the Adelman and Robinson's World Bank model (hereafter AR) is different from the tradition following Johansen is that wages, as well as prices, are made endogenous. But similar with all CGE models, it is built around an input–output table which is extended to include factor substitutability, explicit price modelling and numerous accounting constraints.

AR's outline of their model begins with the production side. Requirements for intermediates in the production process are according to standard technical coefficients and the Leontief production function. As with all input–output modelling, any demand for intermediates is assumed to be met without friction. The demand for labour and capital follows neoclassical profit maximisation, in that each factor is hired until the cost (wages or capital rental rates) equal marginal revenue. These costs are assumed to be identical across all sectors and all factor markets are assumed to clear.

AR then outline the consumer side of the model. Each factor type is represented by a "homogeneous group of people whose consumption behaviour can be represented by an aggregate expenditure function" (AR, p. 22). Expenditure is a function of aggregate

nique was never the most effective method for doing CGE computations. Even those CGE modellers who embraced the Scarf technique in the 1970s had by the 1980s largely abandoned it." (Dixon and Parmenter 1996, p. 7) income and of all prices, requiring a large number of elasticities. The expenditure functions fulfil Walras' Law and are not responsive to changes in overall price level.[47] This property is "is extremely important for proving that the general equilibrium system has at least one solution because the proofs require that one be able to work with the set of normalized price vectors" (DMR, p.498).

Thus the CGE model of Adelman and Robinson (1978) is a pure combination of input–output intermediate-demand modelling and neoclassical economic model where prices, to within a scale factor, are set such that all markets clear. This scale factor is a price index, a simple linear combination of each equilibrium price.[48] Making the crucial difference between this World Bank model and Johansen's model, the AR then develop an iterative solution procedure, adjusting prices gradually until markets clear:

> We have chosen to iterate on prices following a kind of *tâtonnement* procedure.[49] First, start with an initial guess at prices. Second, solve the factor-market equations for wages, employment, and production. Third, solve the product-market equations and calculate excess demands. Fourth, raise or lower prices in sectors where there are excess demands or supplies. Fifth, normalize prices...and start another iteration. Stop iterating when all excess demands equal zero (Adelman and Robinson 1978, p. 24).

AR then extend their model to include inflation (via a supply and demand for money formulation, citing Keynes and Pigou), a great deal of microeconomic detail about households and firms, foreign trade, a government sector (which collects revenue and distributes transfers) and some differentiation between sectors in terms of their ability to exploit monopoly power. Finally, a dynamic version of the model is sketched out where each period's model can be solved individually and is only loosely coupled to the previous year's modelled variables.

Although this model does not claim to produce intertemporally efficient outcomes, nor any clear efficient path (the authors categorise it has having a "lurching equilibrium") they conclude that "we ought to proceed further in the "lurching equilibrium" direction and shift away from the traditional exclusive preoccupation of neoclassical economists with equilibrium models" (Adelman and Robinson 1978, p. 36). The authors are

[47] In technical terms, they are homogeneous of degree zero with respect to prices. Thus $e(P) = e(\lambda P)$ where λ is any constant and P is a vector of prices.

[48] In AR, it is not made clear how exactly this linear combination is constructed. Their equation (12) on p. 23 states that $\bar{p} = \sum_i \omega_i p_i$ but it is never specified what ω_i should be. Based on other methods in very similar models, we assume ω_i is the equilibrium quantity of good i.

[49] From the French for "trial and error," the term was introduced by Walras in an imagined auction in which the auctioneer knows the demand at every possible price and can therefore set the price such that supply is perfectly exhausted.

thus attempting to use the within-period equilibrium assumptions of neo-classical economics and, at the same time, present themselves as representing a break from the intertemporal optimisation of neoclassical models.

Given the status of the authors as World Bank economists, and therefore at the heart of the global establishment, this is an interesting act of economic rebellion, and one which perhaps set the CGE modelling community apart from 'mainstream' economics from the late 1970s on, a split which persists to some degree today.

2.8.5 GTAP

The progress in empirical, computable models by Johansen was not immediately taken up and replicated. In their thorough description of CGE/AGE models and their solution methods, Dixon and Parmenter describe a move to more theoretical work:

> Following Johansen's contribution, there was a surprisingly long pause in the development of CGE modelling with no further significant progress until the 1970s. The 1960s were a period in which leading general-equilibrium economists developed and refined theoretical propositions on the existence, uniqueness, optimality and stability of solutions to general equilibrium models. Rather than being computable (numerical), their models were expressed in general, algebraic terms (Dixon and Parmenter 1996, p. 6).

Furthermore, a schism had developed in the literature between those who solved models using Johansen's linearisation/matrix-inversion approach and those who found numerical solutions to the full non-linear specifications, following Gerald Scarf (Scarf et al. 1967; Adelman and Robinson 1978; Shoven and Whalley 1984).

Realising that modern general-purpose computing had allowed the conversation to move beyond methods of solution and should focus again on economic content, Thomas Hertel, a researcher who was "disillusioned with how CGE modeling was being conducted" appealed to the community to come together, saying that "the two schools of AGE modelling have a great deal in common. Both would benefit from greater interaction" (Hertel, Horridge, and Pearson 1992, p. 400).[50] Under Hertel's leadership, and in a spirit of collaboration and open-sourcing of data and models, a collaboration started in 1993 by the Center for Global Trade Analysis at Purdue University called the Global Trade Analysis Project (GTAP). Hertel wanted to "create a

[50] A description of the conditions which led to Thomas Hertel creating the GTAP project after a scholarship trip to Australia is given at the GTAP website, gtap.agecon.purdue.edu, under the rubric "GTAP history."

platform on which people could conduct consistent and comparable analyses of policy alternatives in global trade"[51]

By the middle of the decade, the group had constructed a CGE/AGE model and was already referring to it interchangeably with the name of the group: "GTAP (Global Trade Analysis Project) is a multi-regional AGE model which captures world economic activity in 57 different industries of 66 regions" (Brockmeier 1996, p. 4). The model is outlined in detail by Hertel and Tsigas (1996) as part of what is called by the project today, "the GTAP book" (Hertel 1997). A commonly-used implementation of GTAP in the numerical-solution package GAMS, is due to Rutherford (2005) and is occasionally cited, erroneously, as the primary reference to GTAP.

We shall, in this short exposition, keep with the project's habit of referring to the model by the name of the project. Following Hertel and Tsigas (1996), we will first outline the accounting identities, and then describe the behaviour of the various actors in the model. All page references are to Hertel and Tsigas (1996) if not specified. All flows in the model are expressed in value terms, not quantity terms, which means great care is taken in the original exposition to track which prices (consumers', producers', foreign, domestic, etc.) a particular flow is measured in. We will not reproduce that careful effort here, as it adds little to the understanding of the workings of the model, but a great deal to the complexity of the exposition.[52]

The principal demand-side actor in the GTAP model is called a "regional household" (Howe 1975). This household has a Cobb-Douglas utility function which distributes income in fixed proportion between the categories of final expenditure: private, government, and savings. The inclusion of a "regional household" actor, rather than simply saying that the three categories of final expenditure remain in fixed proportion to one another, has two stated purposes. The first is to enable the modeller to change the nature of that relationship. The example given is that government purchases and/or savings can be given exogenously, in which case the "regional household" has a budget constraint which dictates private expenditure. The second reason for including a regional household is that it has a utility function, and can therefore arbitrate in welfare terms between e.g. a reduction in government spending and an increase in household spending.

[51] GTAP informational video available at youtu.be/gUo5zin8eDk.

[52] The authors deal only in the briefest way with the issues arising from the combination of working with value flows and employing a neoclassical market clearing model structure, which requires quantities. They simply state that "any market clearing condition can be converted to value terms by multiplying by a common price. In so doing, we circumvent the need to partition value flows into prices and quantities. This has the added benefit of vastly simplifying the problem of model calibration" (Hertel and Tsigas 1996, p. 29).

The regional household earns money by selling its "endowment commodites" to the second actor in the model, the producer.[53] Of these, labour and capital prices are perfectly elastic, whereas land prices are "sluggish" (p. 23). The producer (i.e. firms) combines labour and capital with intermediate goods to produce goods for final demand, which are then sold back to the regional household for distribution among the consumption actors.[54] Firms do not earn any "pure profits," but rather distribute all turnover to other agents in the model.

Imports are determined separately for households, the government, and firms. This is done to allow "import intensities" to vary across uses as was empirically observed. Two more agents are added to deal with international trade. The first is a global bank, which makes up for any shortfalls in domestic demand for saving vs. firms' investment demands. The second is an international transport sector which absorbs the difference between f.o.b. exports and c.i.f imports. With the addition of various taxes, all of which accrue to the households in the region they were levied, the accounting relationships are complete.

Production is done via Leontief production functions for intermediates and CES production functions for deciding the levels of each factor required, and deciding between domestic vs. foreign goods. All production has constant returns to scale. Imports are decided "at the border" meaning that the mix of export origins is the same in each sector.[55] To work out the level of imports required, the firms first determine where they would like their imports to come from, the import mix, and then make a decision about domestic vs. imported goods based on the composite import price implied by their chosen import mix. This is called the "Armington approach" to modelling import demand.

Governments behave identically to firms, with Cobb-Douglas constant budget shares determining how the government's income is spent on goods.

Households have a utility function with a form that was first suggested by Hanoch (1975). In his work, Hanoch complains that existing production/utility functions had

[53] The endowment commodities are land, labour and capital. As usual, capital depreciates proportionally to the size of the capital stock.

[54] Hertel and Tsigas (1996) is peppered with the acronyms of detailed names for each of the concepts in the model, presumably taken from the fact that the model was to be implemented in computer code, a procedure that requires giving mid-length names to each variable (longer than w and shorter than "wage earnings by households"). For example, the flow of intermediate goods, often denoted z_{ij} in input–output studies, is here referred to as "Value of Domestic purchases by Firms at Agents' prices," shortened to VDFA.

[55] This is identical to the way our model deals with imports. In our exposition we talk about a warehouse into which all exports go before being distributed to the various sectors for intermediate use. Domestic sectors are then unable to distinguish between products of different origins.

too many parameters to be estimable. The typical approach at the time was to use CES functional forms, of the type:

$$Y = A\left(\alpha K^{\gamma} + (1-\alpha)L^{\gamma}\right)^{1/\gamma}$$

which are highly non-linear. Estimating these CES forms typically involved taking a second-order Taylor approximation around $\gamma = 0$.[56] These second-order approximations involve large numbers of parameters, scaling with square of the number of factors n. Furthermore, he criticised CES utility functions for implying a constant product mix as the consumer got richer. This is clearly not what the data suggest, since it is well known that wealthier consumers spend a smaller fraction of their income on, say, housing, than do poorer consumers. In other words, real utility functions are not *homothetic*.

Hanoch developed a non-homothetic functional form which allowed for the usual distributional parameters, choosing between goods, but additionally allowed for substitution parameters (rather than the constant substitution elasticity of CES) and expansion parameters which allowed the preferences between goods to change as the budget increased. An additional benefit is that the number of parameters scaled proportional to n not to n^2, thus allowing systems with large numbers of goods to be estimated. This is vital in CGE modelling, where there can be hundreds of sectors.

Instead of elasticities of substitution being independent of total expenditure, the *ratio* of elasticities is independent. In other words, if the elasticity of substitution between goods i and j is denoted as ES_{ij} then:

$$\frac{ES_{ik}}{ES_{jk}}$$

is constant in total expenditure and independent of k. This leads to the name "Constant Ratios of ES" or "CRES." The particular form is estimable using only data on total expenditure (or output, depending on the application), prices and quantities. It is this non-homothetic, estimable functional form which Hertel and Tsigas (1996) use for household demand. The original CRES name is never used in the GTAP literature. Instead it is referred to throughout as constant difference of elasticity (CDE). The CDE functional form is expressed in implicit form in terms of utility U, the price of commodity i, P_i and total expenditure as:

$$1 = \sum_i \beta_i U^{\alpha_i \gamma_i} \left(\frac{P_i}{X}\right)^{\gamma_i}$$

[56] It is well known that the limit of a CES function as γ goes to zero is a Cobb-Douglas function.

were α_i, β_i and γ_i are, respectively, the expansion, distribution and substitution parameters.[57] Thus, this can be viewed as either a utility function, or an expenditure function, depending on which variable is held constant and which is to be solved for.

The task in GTAP of this equation is to distribute the regional household's income between private, government and savings expenditure, with a simple accounting budget constraint:

$$Y = Y_P + Y_G + Y_S$$

McDougall (2003) noticed that this implicit utility function *implies* (although we will not prove it here) that each of the associated expenditure functions—private, government and savings—are homothetic. But as we have seen, the functional form was selected precisely because it was *not* homothetic and so "the demand equations [of Hertel and Tsigas (1996)] are invalid" (McDougall 2003, p. 8).

In order to fix this inconsistency, he proposes a new budget constraint of a more general form, where expenditure shares are allowed explicitly to depend on price. Section 2.3 of McDougall's paper gives complete details. He concludes that "in practice...with standard data bases, errors in the utility results are likely to be small" (McDougall 2003, p. 16).

GTAP remains an active community to this day, with new versions of the database being released every 3 or 4 years since 1993. The latest version of the database includes 140 regions and 57 production sectors. Many of the latest uses of the system centre around energy, climate and the environment but there are some pure trade applications which are still being published. A recent review of GTAP of its use in agricultural and environmental economic analysis were upbeat about its continuing relevance:[58]

> The GTAP project (including the network of researchers) is highly responsive to the policy problems and the tools required for their address...as long as there are policy debates that demand global economic analysis we may expect GTAP to continue its string of success in contributing new knowledge to the policy process (Keeney, Narayanan, and Valenzuela 2017, p. 53).

[57] The GTAP literature differs on how these parameters are denoted. We have here chosen Greek letters to distinguish them from the Latin letters of the variables.

[58] This upbeat review should be read with one eye on the fact that two of the authors, Keeney and Narayanan, work at the Center for Global Trade Analysis, for the GTAP project itself.

2.9 Other approaches

In this section we will cover a number of other approaches to global trade modelling, namely the elasticity approach and the microfoundations approach. This last may seem too large a literature to be included in an "other approaches" section but, as we shall see, the microfoundations literature we survey is largely theoretical in nature and, beyond its significant impact on CGE/GTAP modelling, reviewed in the previous sections, is only tangentially relevant to a literature survey on empirical trade models.

2.9.1 Elasticity and exchange rates

The sheer quantity of trade data available, both the disaggregation into products and the length of the time series, has made econometric studies of trade a rich field of research. In the process of trying to estimate how the volume of trade responds to various economic observables, and following a very wide literature, some of which we will trace out here, Goldstein and Khan (1985) outline a model of import demand which is worth examining in this context. This is principally because it was partly responsible for an increased interest in the role of exchange rates in explaining volumes of international trade.

Their model followed in a tradition referred to variously as the "elasticity approach," the "imperfect substitutes model" and the "Bickerdike–Robinson–Metzler" (BRM) model after the researchers who studied exchange rates in the 1920s (Bickerdike) and 1940s (Robinson and Metzler).[59] Goldstein and Khan (1985) is one of three similar expositions of this approach, along with Dornbusch (1975) and Rose (1991).

The BRM model is very simple and is just as well outlined in words as in mathematics. There are two regions, the home country and the foreign "country," which may be taken to be some aggregate of the rest of the world. Following a devaluation in the home currency, the home price of imports rises, and the foreign price of exports falls. This leads to a decrease in exports and an increase in imports. What effect this has on the balance of trade is ambiguous, since this is measure in *value* terms, not in quantity terms: import quantities fall, but import prices rise. Which effect dominates depends on the price elasticity of demand for imports. The overall terms of trade thus worsen or improve, depending on a combination of the import and export elasticities of both the domestic and the foreign region.[60]

[59] Although the BRM model was, in the early 1990s, described as "somewhat passé" (Rose 1991, p. 301), it was still widely used in the 1970s, enjoying "substantial popularity in policy discussions and interpretations of current events" (Dornbusch 1975, p. 859).

[60] To be precise, we can denote an elasticity as ε, with a subscript of m or x, and an asterisk denoting a foreign variable. Then the terms of trade improve following a devaluation if $\varepsilon_m^*/(\varepsilon_x+\varepsilon_m^*) > \varepsilon_x^*/(\varepsilon_x^* +$

Beyond the primary purpose of investigating the effect of exchange rates on balance of trade, the BRM model hides an important assumption: that it is possible for the price of a domestic good to be different to the price of the equivalent good when imported. This leads to the fact that imports are not perfect substitutes for domestic goods. Goldstein and Khan (1985) outline why this assumption makes both intuitive and empirical sense. If foreign goods were perfect substitutes for domestic goods, there should be no possibility of observing a country both importing and producing a good, and there should be no two-way trade, whereby countries reciprocally trade the "same" product back and forth. As the authors put it, under perfect substitution, we should observe: "(i) either the domestic or foreign good swallowing up the whole market when each is produced under constant [or decreasing] cost... and (ii) each country as an exporter or importer of a traded good but not both" (Goldstein and Khan 1985, p. 1045).

Goldstein and Khan (1985) sketch out an extremely simple version of the imperfect substitutes model in eight equations, the majority of which are accounting identities. The crucial information in the model is containing in their equation 2.1, which states that the level of demand for imports in country i is an increasing function of the importer's income, and that it depends positively on the price paid by importers, and negatively on the price of domestic goods. In a perfect-substitutes world, these simple relationships would be complicated by the fact that the demand for imports would be zero if domestic prices were even one cent lower than import prices.

Additionally, the supply of exports is assumed to depend positively on prices received by the exporters and negatively on domestic prices. In other words, exports are in "competition" with domestic supply. The authors emphasise this fact in criticising the then-existing literature estimating trade flows for assuming that exporters are always happy to supply whatever demand is placed on them, going on to claim that single-equation estimation of price and quantity miss this subtlety:

> An advantage of presenting the supply side as well as the demand side of the imperfect substitutes model is to make it plain that the relationship between quantities and prices is, at least in theory, simultaneous. Despite this fundamental point...the bulk of the time-series work on import and export equations has addressed the supply side only by assumption. [It is unlikely] that even a large single country can increase its total export supply at a constant price unless there exists a large pool of unemployed resources in the export industry itself or elsewhere in the economy (Goldstein and Khan 1985, p. 1048).

ε_m). Thus, an exchange rate devaluation will only help an economy if the quantity effect associated with selling more exports outweighs the cost effect of more expensive imports. This is true when certain conditions on the various elasticities are met, a set of circumstances called the "Marshall–Lerner condition." Much of the empirical trade elasticity literature is concerned with investigating whether and when this condition holds.

The model sketches out the direction of supply and demand levels with respect to income, prices and the exchange rate, and assumes that all markets clear such that supply equals demand. Prices are determined such that this holds true. The importer levies a proportional import tariff, and the exporter receives a proportional subsidy. Quantities and import and export prices are endogenous. Income, overall price levels, tariffs, subsidies, and exchange rate are exogenous. The authors criticise the exogeneity of overall prices, saying that the evidence suggests that they are "strongly influenced" by exchange rates, and that the export supply functions unrealistically assume complete price-taking in assuming that the export price is exogenous to the supply of exports.

They go on to extend this two-region model (country i and the rest of the world) to a multi-region model via elasticities of substitution determining export levels based on relative export prices (adjusted for exchange rate.) And finally they offer some advice on the econometric specification, in terms of lags, simultaneity of the variables, and the stability of the relationships over time before presenting some results and some advice related to the use of currency devaluation as a policy tool.

The thorough nature of this work, in terms of its literature review, its critique of its own limitations, the discussion of econometric practicalities, and the thoughtful policy analysis, have led it to be very widely cited and used as the basis for a large literature on price elasticity more broadly. For example, Fagan, Henry, and Mestre (2005) at the European Central Bank (ECB) use the Goldstein–Kahn model as a basis for the relationship between the Euro area (treated as a single economy) and the rest of the world. This area-wide model (AWM), which the authors first described in a working paper of 2001, is still in active development today by a group including ECB economists.

A widely-cited direct descendent of Goldstein and Khan (1985) builds an econometric model to explain trade among Canada, Germany, Japan, the UK, the US, and OPEC, as well as two other blocks representing the rest of the world. This is primarily done through estimating bilateral price and income elasticities for these regions. Marquez (1990) uses a formulation for import volume which includes own- and other-prices, the incomes of the trading partners in question, and a full set of price lags and dependent variable lags. He additionally adds "potential real income" in the denominator, thus expressing income as a fraction of its potential, allowing for effects of business cycle diversions of observed income from potential. The presence of the lagged left-hand side variable on the right-hand side causes some complication, requiring a complex joint-normality assumption and a distribution generated empirically using Monte Carlo methods. He finds that income elasticities are largely "high," except for Japan and the less-developed countries and all price elasticities are negative, as one would expect, although only half are significantly different from zero. He also finds weak evidence in favour of the Marshall-Lerner condition that a currency devaluation should improve an economy (see footnote on page 69).

In a similar era, Rose (1991) attempts a direct estimate of the Marshall-Lerner condition. In contrast to Goldstein and Khan (1985), Rose's formulation is non-structural. Instead, he postulates a relationship between balance of trade, B, on the one hand, and the exchange rate, q, income, Y, and foreign income, Y^* on the other:

$$B = B(q, Y, Y^*) \qquad \text{(Rose eq. 5)}$$

The functional form is not specified, but this is not a problem since Rose is only concerned with finding the sign of $\mathrm{d}B/\mathrm{d}q$. He can therefore simply estimate a log-linear approximation to his equation (5) with a number of lags. Despite a number of increasingly complex econometric techniques, Rose fails to find any relationship between exchange rates and trade balance, concluding on a gloomy note:

> ...despite the generality of the techniques employed, none has revealed any strong relationship between the variables in question. Loosely speaking, the evidence presented is consistent with the hypothesis that the trade balance can be treated as exogenous with respect to the exchange rate. More precisely, the data do not seem able to reject the hypothesis that the generalized Marshall–Lerner condition does not hold. The negative finding of this paper constitutes a warning to users of the popular partial-equilibrium [i.e. imperfect substitutes] model. (Rose 1991, p. 316)

2.9.2 Trade with microeconomic foundations

All the work we have reviewed so far seeks to explain the levels of global trade as being due to differentials in price, in transport costs, or in income. A different approach to explaining levels of trade (and indeed attempting to explain the existence of trade *per se*), begins with the idea of a "representative consumer" with a utility function and a budget constraint, as per microeconomics. In this branch of the literature, firms are modelled explicitly, entering and exiting the market, and producing goods of a more or less varied nature. Papers which attempt to find 'microfoundations' for macroeconomic equations are not the focus of this literature, but a few very well-known examples have had a big enough impact on the macro-level modelling of trade that they are worth mentioning here.

One crucial idea in explaining trade is to give "the representative consumer" an inherent desire for variety. This simple idea, expressed in elegant mathematics, allows the theoretical economist to explain why countries trade products reciprocally and why countries which produce a product also import it.

The theory begins with the argument of Chamberlin (1950) that, since customers are diverse, they must have diverse preferences. This means that, when taken as a group, they require differentiated products. The fact that producers must differentiate their products, appealing to some potentially narrow group of consumers over others, implies that competition between firms cannot be perfect and that there therefore must exist some degree of monopoly power in the market. Although since these product differentiations are to some degree similar, they must also be "engaged in competition with others nearby on the chain of substitutes." He concludes that both reality, and an ideal economy in terms of welfare, "involves a blend of monopoly and competition and is therefore correctly described as...monopolistic competition" (Chamberlin 1950, p. 86).

This idea is taken up by Dixit and Stiglitz (1977). They argue that leveraging economies of scale allows resources to be saved by producing fewer goods but at greater quantity. Using Chamberlin's diversity-loving idea, this reduction in variety results in a welfare loss to the consumer. This implies that there is a trade-off, in welfare terms, between quantity and diversity. Unlike Chamberlin's work, they make the love of diversity explicit by using a CES utility function in which consumers select between products x_i within a sector, which are good substitutes for one another, and also between all that products of that sector and the rest of the economy, represented by a single numeraire, x_0:

$$U\left(x_0, \left\{\sum_i x_i^{\rho}\right\}^{1/\rho}\right)$$

This formulation allows for both monopolistic competition in production, and both intra- and intersector interactions in demand. Dixit and Stiglitz go on to show theoretically that this diversity and monopoly power is in fact socially optimal, given the utility function.

The first to apply this insight to international trade was Krugman (1980a). Krugman wanted to explain why countries with similar factor endowments and technologies trade with one another. He also wanted to provide some theoretical justification for the empirical fact that countries tend to export those goods for which there is a large domestic market. Using Dixit and Stiglitz's diversity-loving economy and allowing two such economies to trade, he shows, theoretically, that

> if the two countries have sufficiently dissimilar tastes each will specialize in the industry for which it has the larger home market. Obviously, also, each will be a net exporter of the class of goods in which it specializes. Thus the idea that the pattern of exports is determined by the home market is quite nicely confirmed. (Krugman 1980a, p. 957)

Krugman in the same paper also adds transport costs to his model, which allows a large strand of the literature to find microeconomic justification for the gravity model. Well-known examples of this are Bergstrand (1989) and Chaney (2008). Melitz (2003) is another, extremely well-known, descendent of the work of Dixit and Stiglitz and Krugman, and one which contributed to the important extension of the gravity of Helpman, Melitz, and Rubinstein (2008) described in the section beginning on p. 29.

In studying why developing countries have found it persistently difficult to "catch up" to the developed world, Acemoglu and Ventura (2002) develop a very popular trade model which uses the concept of a "continuum of countries." This idea, that there are an infinite number of countries, each of which occupy negligible "mass" along a line of real numbers between zero and one takes its roots from the work of Becker (1952) who attempted to move trade modelling on from using two-countries. This latter work using geometrical argumentation to show that assuming an infinite number of countries allows the use of smooth, neoclassical offer curves.[61] In doing this, he allows a description of "the rest of the world" in a trade model where, for obvious reasons, little else can be specified about how the rest of the world operates as a single economy. This early work seems to have had little impact on the trade modelling world until much later.[62] The review of classical international trade theory by Chipman (1965) discusses Becker but does not mention his use of an infinite number of countries. It is not until a 1988 work on economic development which uses the continuum idea in the context of developing countries since there are many of them and, economically, they tend to be small: "The simplest kind of world to think about is one with perfectly free trade in the two final goods and with a continuum of small countries, since in that case prices in all countries will equal world prices...and each country will take [them] as given." (Lucas 1988, p. 31)

[61] Offer curves are an idea developed by two of the founding fathers of modern, mathematical economics, Marshall and Edgeworth. They describe the amount a country needs to import of one good in order to be able to export each quantity of another good.

[62] Although trade models using a continuum of countries did not take off until the late 1980s, models with a continuum of goods were already common by the time of the review of the effect of technology on trade by Grossman and Helpman (1995). They give Dornbusch, Fischer, and Samuelson (1977) as an example, although that paper cites many earlier works. Because this strand of literature exclusively studies the two-country case, it is outside the scope of this review.

The Acemoglu and Ventura (2002) model extends the modelling technique of using continua to model large collections: there is a continuum of countries *and* a continuum of intermediate goods, as well as there being infinitely many firms in each country. This technique allows them to use integrals throughout, where normal counting models would be forced to use sums. This in turn makes it natural for the authors to talk about "a law of motion for the world economy." which they then use to analyse the dynamics of terms of trade, rate of return to capital and rate of capital accumulation.

This is clearly a powerful approach. But we think that this takes us about as far as we can go from the empirical, data-driven approaches reviewed in the previous sections. It is therefore here that we conclude this review of the literature on global trade modelling, while acknowledging that no review can ever hope to call itself complete.

Part II

The model

Chapter 3

A simple demonstration model combining trade and input-output

3.1 Introduction

In chapter 1 we described a model which combines input-output tables as country descriptions with an international trade model, both of which were based as far as possible on data. Here, we present that model. We will refer to it throughout this thesis either generically with terms such as "our model", or by the name "the demonstration model".

There are two versions of the model which are used throughout this thesis in each of the chapters. The first is a 40-country model based on data from 2010 and the second is a version extended by the estimation process we will outline in chapter 5 on p. 138 which is based on data from 2005 and contains closer to 200 countries. At the start of each relevant chapter we will state, in a footnote marked with an asterisk, whether the analysis was performed with the 40-country model or the 200-country model.

We begin by providing a short overview of the mechanics of the demonstration model and some words on the motivation for the modelling choices we made in designing the model. A huge amount of detail on the data sources referenced throughout this chapter, and some of the decisions made about how to handle and manipulate that data is given in Appendices B to D.

3.1.1 An overview of the concept

The demonstration model is a representation of countries, their economic structures and the trade relationships between them. It is constructed in two distinct but interdependent parts, one at the country level and one at the international level. The first represents the internal workings of each country's economy, and the second describes each country's interactions with the rest of the global economy.

The national part of the model is based on input-output (IO) tables with 35 sectors taken from the World Input-Output Database (WIOD).[1] See section 3.3.1 on p.88 for detail on how these input-output tables are used. The international part uses trade relationships found in the UN's commodities trade database, COMTRADE, and its services trade database, ServiceTrade.

The IO tables are modified to reduce the number of separately specified input sectors. Where the WIOD tables include two input rows for every sector, one for domestic use and one for imports, the demonstration model combines these into a single row. To do this, we assume that domestically produced goods and imported goods are available in fixed proportions and are indistinguishable from one another from the perspective of the domestic sectors' input requirements. We call these fixed proportions "import ratios". More detail on these can be found in section 3.3.2.

The national IO models are connected together in the second part of the model. Countries are assumed to import each sector from each other country in a fixed pattern. This means that the ratios in which a given country imports a given product from each other country remains fixed. These fixed proportions are calibrated using the UN trade data mentioned above. The proportions are called "import propensities" and are discussed in detail in section 3.3.3.

3.1.2 Motivation for the model design

Broadly, the aim when designing the demonstration model was as a happy medium between on the one hand the sometimes ferocious complexity of CGE/GTAP models, with non-linearities sometimes nested three deep and working descriptions of households, firms and government sectors, and on the other hand the simplicity and linearity of input-output, with its elegant, but theoretically restrictive, matrix inversion solution method.

The structure outlined in section 3.1.1 can, in comparison to the mathematical elegance of input-output, appear overly complicated. Multi-region input-output modelling (MRIO) already provides a method to combine national input-output models into a global system, by simply extending the list of sectors to include a distinct sector for every country-sector pair. In a national IO table, sectors are arranged in columns and each column has associated with it a row value for each other sector. The value recorded at

[1] See Appendix C and particularly table C.2 on p.312 for a complete list of the 35 sectors as well as the short names which will be used throughout this thesis for convenient reference.

the intersection between a row sector and a column sector represents the intermediate flow from the row to the column.

In our case for example there are 35 sectors, so the IO table is 35×35 (ignoring imports and other complications). With MRIO, each sector has an associated row for each sector *in each country*. Thus, in our case with 40 countries, an MRIO table would have $35 \times 40 = 1{,}400$ rows and columns. The great strength of IO modelling is that it is empirically derived. Indeed the IO table is really just a convenient way to store data gleaned from interviewing businesses about their input costs. Because of this, the MRIO table we have described would imply that 1,400 pieces of information have been captured empirically, recording flows from every country/sector to every other country/sector in the model.

This is clearly not the case; many (or all) of the off-diagonal entries in the table (i.e. those relating to international relationships) will have been derived in some way from trade relationships. But the MRIO table makes no distinction between data points and modelled values, thus losing the appeal of IO as a data-driven empirical methodology.

Our setup clearly distinguishes between empirical values (the intermediate IO flows and the UN trade flows) and the modelled coefficients which result from them (the import ratios and import propensities.)

Furthermore, if each country we add to the model requires 1,400 pieces of information for every sector in that country, we are greatly restricted as to the number of countries we can include. Since the ambition of the demonstration model is to be a global model, it is important that the number of pieces of information required per country be as small as possible. Per country, our setup requires $35 \times 35 = 1{,}225$ intermediate flows and 35 import ratios to be calculated or estimated for each country added to the model, instead of the $35 \times 35 \times 40 = 49{,}000$ for MRIO. This orders-of-magnitude difference in data requirements has important implications for the estimation of countries who publish little or no data on their economic structure. This is a problem we address in chapter 5.

Finally we provide some justification for the separation of the national part of the model and the international part of the model outlined in section 3.1.1. Input-output has rightly remained controversial among many economists due to its use of the linear production function and lack of discussion of substitution and demand elasticities (Oosterhaven 1989; Oosterhaven 2012). Extending the IO table to an MRIO can only serve to exacerbate these complaints, particularly the latter. By separating the internal economic structure from the global trade part of the model, we are allowing for future developments which model either part or both parts in a more 'interesting' way, without requiring a single modelling solution for the entire global economy such as that provided by MRIO.

Users of the model who are more comfortable with the complexity and assumptions required by CGE/GTAP modelling, for example, may like to incrementally add the nonlinearities which are the hallmark of those approaches. The decoupling of these parts of the model allow for production and trade to work according to different mathematical specifications, something which MRIO inherently cannot do.

The model structure we present here owes much to the early iterations of Project LINK, discussed at length in section 2.7.3. It is a combination of the Mini-LINK, with its iterative solution, and the Maxi-LINK in the way that imports are first determined from country models, and exports are then determined to meet import demand. Since the focus of this thesis is on the applications of global modelling, rather than the model structure itself, the spirit of Project LINK, which aimed to provide a simple, quick model on which to build, suits us well here.

The work of this thesis makes particular use of the World Input-Output Database (WIOD), due to Timmer, Dietzenbacher, et al. (2015). The IO tables assembled as part of the WIOD project were created during a resurgence of interest in such global input-output data. Indeed, Tukker and Dietzenbacher describe the study of global multi-region input-output (GMRIO) models as having "…only just begun. An exciting future lies ahead for the field of GMRIO." (Tukker and Dietzenbacher 2013). Three other major global examples of input-output databases were released within a short space of time: Exiopol (Tukker, Koning, et al. 2013), GTAP (Walmsley, Aguiar, and Narayanan 2012) and EORA (Lenzen, Moran, et al. 2013).

The WIOD was chosen above these alternatives because of the rigour of its estimation procedures; the completeness of its documentation, including a method, vital for the present research agenda, for linking product trade information from the UN back to the input-output tables; its inclusion of tables in current year and previous year prices which allows for proper price-independent time series analyses; and lastly because it is freely available for download[2]. See p.298 in Appendix A for a fuller discussion of the alternatives.

[2] from www.wiod.org/database/iot.htm

3.2 Commodities, services and sectors

As we have seen, the global trade part of the model uses data on commodities and services trade from the UN. In order to link these figures to the input-output models within countries, we must make a mapping from the UN product categories to WIOD sectors. This section describes the detail, and some of the difficulties, of this process.

Much of the material of this section is included for the completeness of this thesis as a research document and the casual reader is encouraged to skip directly to the formal description of the model given in section 3.3 on p.88.

3.2.1 Assigning commodities to sectors

The commodities flows in UN COMTRADE are stored against a Harmonized System (HS) code, a six-figure hierarchical categorisation. The creators of WIOD produced a mapping from these six-figure codes to 16 commodities sectors (plus a very small number of flows to two services sectors: "Business Services" and "Utilities".) Figure 3.1 shows the total value of all goods flows assigned to each sector in 2010. Several sectors saw global trade of over a trillion $US, the biggest three being Electricals, Mining and Chemicals.

There is a small flow to the "Utilities" sector, made up of two HS codes: "Coal gas, water gas, etc. (not gaseous hydrocarbons)" and "Electrical energy". There is also an even smaller flow to the "Business Services" sector, made up of a single HS code: "Plans and drawings for architectural etc use". It is not clear whether these are well categorised or not, but in both cases the mappings have been kept, leading to a small quantity of COMTRADE flows being categorised as part of services sectors. This is very much the minority.

3.2.2 Commodities which are not assigned to sectors

A far more serious problem than that of commodity sectors categorised as services, is that of uncategorised commodities.

In 2010, the trade in these amounted to almost $US 2 trillion. Section 3.2.2 shows the top 10 largest HS codes by total global trade flow value. At least some comfort is provided by the fact that the largest single HS code is indeed a non-categorical catch-all "Commodities not specified according to kind". This means that a large amount of global commodities trade goes unmeasured whenever we attempt disaggregate by commodity, however categorised.

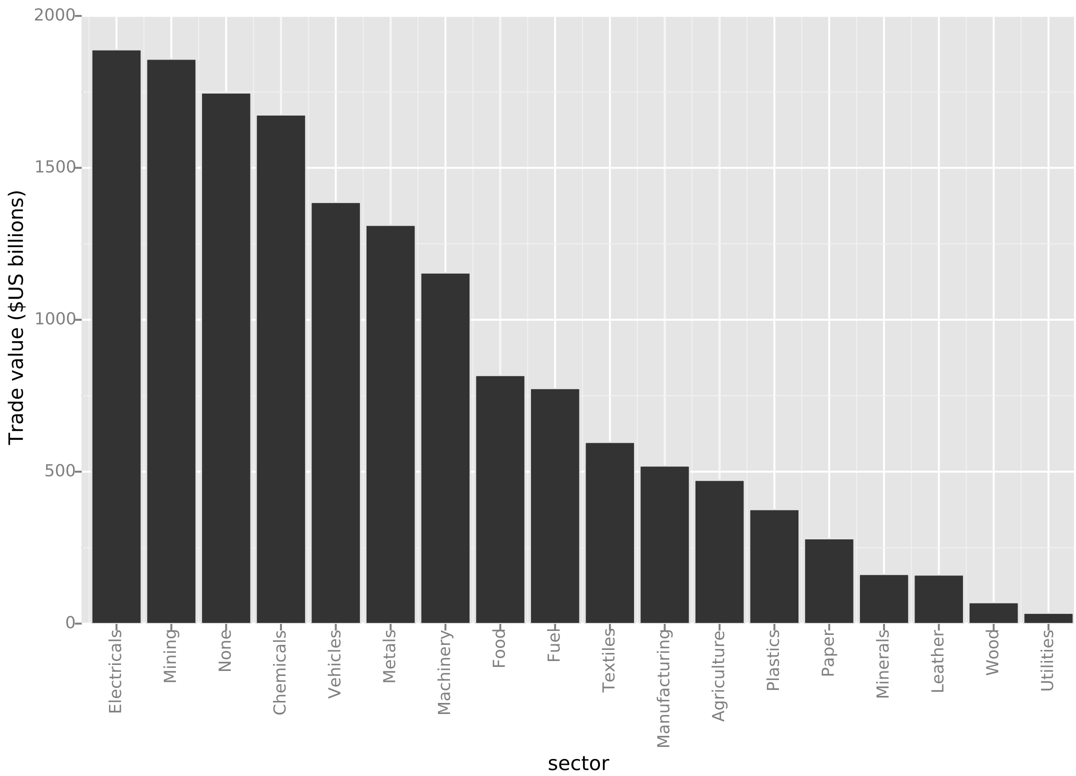


Figure 3.1: UN COMTRADE flows assigned to each WIOD sector in 2010. A very small flow in the "Business Services" sector is not shown.

Table 3.1: The top 10 largest HS codes by total global flow value which are not categorised into a WIOD sector. Since the descriptions are sometimes very long, only the first 50 characters are shown.

HS Code	Description	Trade Value ($US billions)
999999	Commodities not specified according to kind	391
854231	Electronic integrated circuits, processors & ...	191
854239	Other Electronic integrated circuits, other t...	159
851712	Telephones for cellular networks/for other wir...	108
851770	Parts of telephone sets, incl. telephones for ...	83
854232	Electronic integrated circuits, memories	82
851762	Machines for the reception, conversion & trans...	76
844399	Other parts & accessories for printing machine...	55
852872	Other colour reception apparatus for televisio...	52
852580	Television cameras, digital cameras & video c...	35

Table 3.2: The top ten biggest exports of HS 854231, an electronics product not categorised into a WIOD sector.

Exporter	Export Value $US billions
Regional/Unknown	28.81
Malaysia	28.78
China	24.48
Republic of Korea	19.74
USA	17.69
Singapore	16.47
Japan	13.71
Philippines	8.71
Costa Rica	5.10
Germany	4.28

The next five items in the table are all related to electronics, including almost $US 200 billion of trade in "854231 – Electronic integrated circuits". This category is clearly an extremely important global commodity.

It is not clear why this commodity code remained uncategorised in the WIOD correspondence table, since several other 8542 codes are correctly categorised under the "Electricals" sector. One possible explanation is that these codes were introduced after the WIOD categorisation was completed. This is not entirely implausible, since codes are added frequently. The absence of this commodity category in the resulting model will affect countries differently, leading to systematic underestimation of the importance of the "Electricals" sector.

In table 3.2, the ten largest exporters of HS 854231 are shown, in order of the magnitude of their export. The largest single category is a catch-all "regional/unknown", for categorising flows which are either not recorded as being from nation states, but rather from regions, or whose origin is not recorded at all.

It is interesting that Malaysia is the world's largest exporter of this particular commodity code. A report by Credit Suisse (Wan 2013) shows that Malaysia's competitiveness in the early 2000s was driven in part by "printed circuit board assembly and sound cards." This might explain Malaysia's top ranking, above even China and Korea.

Of the countries in table 3.2, China, Korea, the USA, Japan and Germany are all in the WIOD list of 40 countries, and will therefore be underestimated in terms of their Electronics industry exports. Although re-categorising these flows would be relatively straightforward, our main focus here is not present the perfect working model, but to present a schema for how such a model can help give a global perspective on the issues

Table 3.3: Trade totals of the five services BoP codes not assigned to a WIOD sector, in billions of dollars.

BoP Code	Description	Trade total ($ billions) 2005	2010
983	Services not allocated	3,904	773
391	Workers' remittances	104	171
500	Direct investment	65	102
310	Compensation of employees	15	29
431	Migrants' transfers	8	13

we cover in this thesis. This situation has therefore not been resolved in the current iteration of the model, and this should be borne in mind in the analysis which follows.

3.2.3 Taiwan and the commodities data

Taiwan is not included in the trade data for political reasons[3] The WIOD have dealt with this as outlined in the section 'country-specific problems' of the main working paper (*The World Input-Output Database (WIOD): Contents, Sources and Methods* 2012). In summary, the authors took import data from COMTRADE flows categorised as 'Other Asia, Nes' (where "Nes" is "not otherwise specified") and took on board the UN assumption that this only represents flows relating to Taiwan. Additionally, they took export data from the OECD, which does report Taiwanese imports and exports explicitly.

We have not replicated this work in producing the global trade aspect of our model, and therefore Taiwan has a functioning input-output model as its country representation, but no imports or exports to the rest of the world. Therefore, in much of the analysis which follows, Taiwan has either explicitly been excluded, or should be excluded by the reader where the results for Taiwan do appear.

3.2.4 Assigning services to sectors

The mapping from commodity HS codes to sectors discussed above does not include the 19 WIOD services sectors. Services are currently far more coarsely categorised than commodities, and the task of assigning these broad services categories to sectors is more of an art than a science.

The categorisation used by the UN's ServiceTrade database is based on the services sections of the IMF's Balance of Payments (BoP) coding (International Monetary Fund

[3] For more details on this, read the UN description of the situation at http://bit.ly/1oqhuFK.

2009). In the documentation around the construction of WIOD, *The World Input-Output Database (WIOD): Contents, Sources and Methods* (2012, p. 35) describe the process of mapping these services BoP codes to a eurostat (2015) product classification called the Statistical Classification of Products by Activity (CPA). The resultant mapping is non-unique in the sense that a single BoP code is mapped to several CPA codes. By inspection of supply-use tables provided by WIOD, we have mapped CPA codes to WIOD sectors. This mapping is somewhat heuristic and represents at best a good approximation. Where a single BoP code is mapped to CPA codes which span different WIOD sectors (as happens frequently), the trade flows relating to this BoP code have been distributed evenly among the relevant WIOD sectors.

Finally, some BoP service codes could not be assigned to a WIOD sector, and these have been excluded from all the analysis in this thesis. A summary of these services codes is given in table 3.3. Particularly worthy of note is that remittances, money sent to a home country by workers abroad), amounted to $170 billion in 2010. Some countries will be more affected by the omission of these not-insignificant flows than others. As in previous sections, we appeal to the demonstration nature of the model we are building in justifying our decision not to focus more on data issues of this kind.

3.2.5 Country code inconsistencies

Some of the countries in the services trade data are different to those used in the COMTRADE data. We have therefore manually updated the database fields `reporter_code`, `from_code` and `to_code` in the database table `services_flows` to match the country codes used in the goods data, according to the mappings given in table 3.4. Worthy of note is the fact that three country codes, those for the US Virgin Islands, Puerto Rico and the USA have been combined into a single country code.

Many flows are listed against regions (such as Asia, Central and Southern Africa, etc.) rather than against countries. No attempt has been made to distinguish between these regions, and they are all categorised as country code -1, "unknown". Since setting many trade flows to `country_code` -1 either in `from_code` or `to_code` leads to many rows with the same combination of `trade_year`, `from_code`, `to_code` and `service_id`, the `trade_value` column is summed by these four columns before being used.

Services flows recorded against the entity 'World' are sometimes treated as a simple sum total (see Luxembourg 442, 2010, BoP Code 229) and sometimes as records in their own right. To cope with this fact, the database checks that the 'World' and 'non-

Table 3.4: Replacements made to UN services country codes and names in the UN ServiceTrade data, to ensure the data corresponds to that in UN COMTRADE. Note that three codes are combined into the single code 842.

Original		Updated	
Code	Name	Code	Name
250	France	251	France, Monaco
380	Italy	381	Italy
438	Liechtenstein	757	Switzerland, Liechtenstein
850	US Virgin Isds	842	USA, Puerto Rico and US Virgin Islands
630	Puerto Rico		
841	USA		
833	Isle of Man	826	United Kingdom

World' flows don't sum to one another before setting a flag to include each row in any output.

Additionally, BoP Code 200, "Total EBOPS Service" is sometimes included as a total column and other times is the only record for a particular from/to country pair. Records related to this BoP Code are therefore retained only when the latter is true and the BoP Code 200 record is the only one for that from/to pair, regardless of reporter. This last point is a subtlety to take account of the fact that sometimes a BoP 200 flow is the only one reported by a recipient country (and would therefore be included) but the exporting country reports at a finer grain of BoP code, in which case the single recipient-reported record can be ignored.

Where there is mirrored reporting, i.e. where both the importer and the exporter have reported a flow, only the importer's view is retained. This is identical to the equivalent situation with commodity flows.

3.2.6 Questionable data values

Some services flows are recorded as negative. See, for example, Cape Verde (132) Insurance Services (253) imports from the World (0). It is not clear what this is intended to mean, and it leads to all negative flows for that sector in the balancing procedure which follows (see section 3.2.7). All negative flows are therefore treated as zero.

Some services flows at first appear to be questionable in their order of magnitude. A good example of this is BoP 253, "Insurance Services", between the US and Bermuda[4]. The between 2000 and 2011 (and possibly other years) the import of this service into

[4] See the search results at http://bit.ly/1qainE2. Note that a free login to the COMTRADE system is required.

the US from Bermuda was over $1 billion per year. Although the fact that this is relatively consistent over time lessens the probability of an inputting error. It may be that Bermuda is an used by American insurance firms for tax purposes and the profits are therefore booked offshore, leading to possible overstatements in the trade figures for Bermuda. This may also be an issue in other tax-related off-shore schemes.

3.2.7 From services totals to services flows

Unlike the commodities data, many of the flows in UN ServiceTrade are only given as total imports or exports (often recorded as a trading partner of "World".) For the purposes of the demonstration model we need point-to-point flows, not flow totals.

A great deal of effort has gone into estimating point-to-point services flows from flow totals, and this is covered in detail in chapter 4. For now, it suffices to say that many of the services flows used in the calibration of the model which follows have been estimated from totals and are therefore somewhat less reliable than the commodity flows, almost all of which are recorded as point-to-point.

3.3 Formal description of the model

As described in the overview in section 3.1.1, the demonstration model is a collection of coefficients, calculated from data, which describe the economy of each country and its trading relationships with the rest of the world. The derivation of these from data is described in sections 3.3.1 to 3.3.3.

A collection of coefficients becomes a model only through its use in analysing a scenario. In other words, the coefficients must be used to calculate an output relating to some circumstance other than that found in the data. This requires the addition of theory to the collection of coefficients. The demonstration model uses input-output and trade theory to calculate production levels and trade flows of each country/sector for a given set of final demand vectors which are exogenously given. The theory behind this is described in sections 3.4 and 3.5. This section then concludes with a discussion of trade flows to and from countries not explicitly featured in the model in section 3.6.

Throughout all that follows, we will use the following notational conventions. Each coefficient may have up to four subscripts, depending on context. First it may have either one or two country subscripts, denoted i and j, implying a natural direction "from" i and "to" j, the precise meaning of which will depend on the particular context. Second, it may have either zero, one or two sector subscripts, denoted r and s, again implying a direction "from" r and "to" s. Finally, all values which are taken from data will have a bar like so: $\bar{x}$.

3.3.1 Technical coefficients

Each sector in each country (hereafter "each CS", to be read as "each country-sector") buys the output of each other sector which is used as input to the production process. Demand of this kind is called "intermediate demand". Specifically, the intermediate demand placed on sector r by sector s in country i is denoted z_{irs}.

In fact, in WIOD, each z is recorded as a demand on a domestic sector and a demand on a foreign, imported sector (regardless of origin). We will denote intermediate demand on domestic sectors as $\bar{z}_{irs}$ (recall that a barred variable denotes that it is taken directly from data) and intermediate demand on import sectors as $\bar{z}^*_{irs}$. We therefore have the total intermediate demand calculated as:

$$z_{irs} = \bar{z}_{irs} + \bar{z}^*_{irs} \tag{3.1}$$

Given the total production of sector s in country i, denoted as x_{is}, we can calculate the intermediate demand required *per unit* of total production, or output. This quantity answers the question "how much of r is required to make a single unit of s?" and is called a *technical coefficient* and is denoted a_{rs}. It is calculated as:

$$a_{irs} = \frac{z_{irs}}{\bar{x}_{is}} \tag{3.2}$$

Notice that $\bar{x}_{is}$ has a bar. Although it is given in WIOD, it is simply the sum of a number of other data values and is therefore not strictly an observed quantity itself, but it will nevertheless be convenient to distinguish it from calculated total production once the model is solved.

It is in fact calculated using a simple accounting identity which states that production is equal to use. This has the important corollary that no production is stored or wasted: each unit of production goes either to fulfil intermediate demand, $\bar{z}$, final demand of consumers and government, $\bar{f}$, or export demand $\bar{e}$.

Therefore, in a model with S sectors, we have:

$$\bar{x}_{is} = \left(\sum_{u=1}^{S} \bar{z}_{isu} \right) + \bar{f}_{is} + \bar{e}_{is} \tag{3.3}$$

It is the value of $\bar{x}_{is}$ calculated in equation (3.3) which is used in equation (3.2) to calculate the technical coefficients.

This derivation is very similar to that given in (Miller and Blair 2009) but care should be taken over the source of the various quantities. Note that equation (3.2) does *not* feature the data quantity $\bar{z}_{irs}$, but rather a combined domestic/imported quantity taken from equation (3.1). This means that the technical coefficients we calculate in equation (3.2) can be considered as "combined" technical coefficients, describing the quantity of sector r required to produce a unit of s, regardless of the provenance of r (domestic or imported).

Once these a have been calculated, they will be used to calculate total intermediate demand and, hence, total production, from a given level of final demand by a standard technique of input-output modelling given in section 3.4 where the method of solving the model is given.

3.3.2 Import ratios

Since the technical coefficients we have specified in section 3.3.1 have nothing to say about the quantity of imported goods required for a given level of domestic production, we need a method of retrieving import requirements. This requires us to make an assumption about the relationship between imports and domestic production.

Inspired by the description given by Duchin (2004) of Leontief's proposed global model, we treat foreign and domestic products as perfect substitutes. This means that, from the perspective of the agent demanding the good, they are indistinguishable. This assumption will turn out to be important for the goal of coefficient parsimony described in section 3.1.2 on the motivation for the design of the model. [5]

Leontief assumed that engineers in an importing country do not care where a product originated; they will simply know that domestic production does not meet their demand, and instead demand a perfectly-substitutable imported product. In a similar spirit, we assume that when a product arrives at the shores of an importing country, it enters a theoretical 'national warehouse' along with domestically produced products, at which point the two become indistinguishable. Miller and Blair (1985) refer to this concept as *import similarity*.[6]

We assume that the goods will be imported into this warehouse such that the proportion of domestic to imported goods always remains constant. This relationship is assume to be fixed per country and per sector, and is called the *import ratio*. It is calculated, from the data, as the share of imports in total domestic supply. We say total *domestic* supply because we assume that all export demand is fulfilled by domestically produced goods: imports may not fulfil any export demand.

If we denote the sector-s import ratio of country i as d_{is}, we therefore have:

$$d_{is} = \frac{\bar{m}_{is}}{(\bar{x}_{is} - \bar{e}_{is}) + \bar{m}_{is}} \tag{3.4}$$

The term $(\bar{x}_{is} - \bar{e}_{is})$ represents the total domestic supply, which, as can be seen from equation (3.3) is equal to $\left(\sum_{u=1}^{S} \bar{z}_{isu}\right) + \bar{f}_{is}$.

[5] This particular aspect of the modelling approach is perhaps the hardest to justify. WIOD contains detailed data on individual sectors' requirements for foreign versus domestic goods, which is essentially being thrown away here. The justification in terms of coefficient parsimony given here only applies to countries being added to the model through mathematical estimation, not to those countries for which we have data via WIOD. We can perhaps appeal to a simplicity of exposition argument here, by taking an approach which will allow us to treat WIOD countries and estimated countries identically. In future iterations of the model, this may be relaxed and WIOD countries allowed to behave differently to estimated countries *vis-à-vis* their import demand.

[6] An alternative to this would be the very widely employed "Armington approach" which models imports from different countries as imperfect substitutes for one another. This is covered in detail in section 2.6.2 of the literature review. Applying this approach would be an obvious first step towards going beyond the linear nature of this model, and towards the multiple nonlinearities of CGE/GTAP modelling and is a good first avenue for further work.

The import of sector s into country i, $\bar{m}_{is}$, is calculated from WIOD data as what we might term "total imported supply":

$$\bar{m}_{is} = \left(\sum_{u=1}^{S} \bar{z}^{*}_{isu} \right) + \bar{f}^{*}_{is} \tag{3.5}$$

where $\bar{f}^{*}_{is}$ is the final demand for imported goods and is taken directly from WIOD data, as shown by the bar on the variable name. Note that $\bar{m}_{is}$ will *not* in general be the same figure recorded by COMTRADE for the quantity of good s imported by country i.[7]

Once these d have been calculated from data, they can be used to retrieve a country's import demand from its total production when the model is being solved. Section 3.4 on solving the model gives complete details.

We close this subsection with a word about how this assumption helps us achieve model parsimony: the standard open economy input-output models with S sectors require $2 \times S$ technical coefficients for each sector. S domestic coefficients and another S for imports. Our assumption of fixed import ratios halves the number of technical coefficients that must be specified for a given country. The cost of the assumption is that we lose the ability to model the idea that the domestic sectors can have different preferences for imported over domestic goods. This decision about product mix is now made at a national, rather than at a sector, level.

The benefit, as alluded to above, is that we can treat in an identical fashion WIOD countries and countries which are to be added to the model through mathematical estimation. This allows for a greater simplicity of exposition (and, indeed, of implementation).

3.3.3 Import propensities

Trade in each sector takes place between countries according to each country's import demand for that sector. It is one of the strengths of the demonstration model that it separates the determination of import demand from the determination of the international trade network which the set of import demands induces. But in this iteration of the demonstration model the resulting network of trade is determined in an extremely simple way.

The determinants of global trade are complex, and an enormous literature is dedicated to regression analysis seeking to determine which observable variables contribute to or diminish trade. Attempting to model the determinants of observed trade flows is a

[7] We use m for imports instead of i to avoid confusion with the identity matrix I when matrix forms are discussed in section 3.4.

difficult task in a context where data sets are naturally constrained in size (there are only around 200 countries to work with) and where there is no possibility of experimentation or examination of a counterfactual.

Markusen (2013) has a good recent review of that work whose authors have not been intimidated by these restrictions, and gives an overview of the various determinants tested in the literature. Variables recently tested include distance (Egger and Pfaffermayr 2004), currency unions (Bangake and Eggoh 2009), common languages (Melitz 2008) and ethnic links through migration (Egger, Ehrlich, and Nelson 2012). See the literature review in chapter 2 for a full discussion of the history of trade modelling.

While acknowledging that each of these almost certainly *does* influence the level of trade, the approach we take here takes the view that all the complexities of relationships between countries are essentially unknowable, and certainly non-linear, but are encoded into levels of trade we see in the existing observed trade network.

We therefore take the trade data seriously and assume that the state of the trade network for a particular year encodes all the unknowable myriad factors which promote or inhibit trade between a given pair of countries. Countries are assumed to trade with one another in fixed proportions, specific to a sector. These proportions are derived from data, as with the other quantities discussed so far.

Specifically, country j receives a fixed fraction, p_{ijs}, of its total import demand of sector s from exporting country i. We call these fractions *import propensities*, and they are calculated as:

$$p_{ijs} = \frac{\bar{y}_{ijs}}{\sum_{k=1}^{C} \bar{y}_{kjs}} \tag{3.6}$$

where $\bar{y}_{ijs}$ is the COMTRADE flow of sector s from country i to country j. Note that the denominator term in equation (3.6) amounts to a total import of s for country j, but that this quantity will not in general be the same as the WIOD-reported import $\bar{m}_{js}$ used in equation (3.4) in calculating the import ratios.

There are two small caveats with this procedure. Firstly, COMTRADE records some negative flows, with no mention in the documentation what these are intended to represent. These are all set to zero before the calculation in equation (3.6) is carried out. Secondly, no country is allowed to supply *any* of its own import demand. Thus all flows y_{iis} (of which COMTRADE records some) are also set to zero. There is potentially meaning in such flows, which are designated by COMTRADE as “re-exports” and “re-imports” but, in the spirit of our model being a demonstration one only, we ignore any such intended meaning throughout this thesis. Future work may wish to re-examine the role of re-exports and re-imports in the global trade network.

3.4 Solving the model

In sections 3.3.1 to 3.3.3 we have derived a set of coefficients directly from data (both WIOD and COMTRADE) which will allow us to build a model which produces a complete set of productions, exports, imports and trade flows for any exogenously given level of final demand in each country in each sector.

We will take each of these quantities in turn and show in detail how they are to be calculated. We begin with the derivation of total production from final demand in a given country i.

3.4.1 Solving for total production

The derivations in this section are almost identical to the standard input-output equations described in great detail by Miller and Blair (2009). We assume a linear production function as per the original input-output treatment given by (Leontief 1936) [8]. Thus, if a_s units of a particular input s are required to produce a single unit of sector r, then $2a_s$ units are required to produce two units of r and so on.

We also assume that the total amount of sector r available, via imports and domestic production, is precisely equal to total demand; that no goods are stored from one year to the next, and no demand goes unfulfilled. We can split "total demand" into three parts: demand for use as an intermediate in the production process of a sector, "intermediate demand", demand for consumption by households and governments, "final demand", and demand from other countries for whatever use, "export demand".

Intermediate demand for sector r in country i is given by the total productions of *every* sector of country i and the relevant technical coefficients describing how much of r is required to produce each of those other sectors. Thus:

$$\text{intermediate demand}_{ir} = \sum_{s=1}^{S} a_{irs} x_{is} \tag{3.7}$$

We can therefore write:

$$(m_{ir} + x_{ir}) = \sum_{s=1}^{S} a_{irs} x_{is} + f_{ir} + e_{ir} \tag{3.8}$$

[8] Controversy over the many functional forms suitable for use as a production function go back to at least Shaikh (1974) and continues broadly unabated (Dosi et al. 2014). In this context, a linear production function, requiring the fewest possible parameters to be estimated, seems like a not-unreasonable choice. Any attempt at non-linearity with regards to production *must* give some details of where on a particular non-linear curve an economy finds itself. This is very seldom done.

where the left hand side is the total amount of r available in country i through both domestic production and imports.

Due to our assumptions that goods are available in a fixed ratio between domestic and imported, and that imported goods may not supply export demand, we can write an expression similar to that given in equation (3.4) in the section on import ratios (section 3.3.2), but with the ratio now already known:

$$d_{ir} = \frac{m_{ir}}{(x_{ir} - e_{ir}) + m_{ir}} \tag{3.9}$$

$$\implies m_{ir} = \frac{d_{ir}}{1 - d_{ir}}(x_{ir} - e_{ir}) \tag{3.10}$$

$$\implies m_{ir} + x_{ir} = \frac{1}{1 - d_{ir}}(x_{ir} - d_{ir}e_{ir}) \tag{3.11}$$

which we can substitute back into equation (3.8) to give:

$$\frac{1}{1 - d_{ir}}(x_{ir} - d_{ir}e_{ir}) = \sum_{s=1}^{S} a_{irs}x_{is} + f_{ir} + e_{ir} \tag{3.12}$$

$$\implies x_{ir} - (1 - d_{ir})\sum_{s=1}^{S} a_{irs}x_{is} = (1 - d_{ir})f_{ir} + e_{ir} \tag{3.13}$$

Equation (3.13) represents a system of S equations for solving the total production of every sector in country i. For convenience, we can gather the terms into matrices and vectors as follows:

$$\mathbf{x}_i = \begin{bmatrix} x_{i1} \\ \vdots \\ x_{iS} \end{bmatrix}$$

$$\mathbf{A}_i = \begin{bmatrix} a_{i11} & \dots & a_{i1S} \\ \vdots & \ddots & \vdots \\ a_{i1S} & \dots & a_{iSS} \end{bmatrix} \tag{3.14}$$

$$\hat{\mathbf{d}}_i = \begin{bmatrix} d_{i1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & d_{iS} \end{bmatrix}$$

allowing us to rewrite equation (3.13) as

$$\begin{aligned} \mathbf{x}_i - (\mathbf{I} - \hat{\mathbf{d}}_i)\mathbf{A}_i\mathbf{x}_i &= (\mathbf{I} - \hat{\mathbf{d}}_i)\mathbf{f}_i + \mathbf{e}_i \\ \implies (\mathbf{I} - (\mathbf{I} - \hat{\mathbf{d}}_i)\mathbf{A}_i)\mathbf{x}_i &= (\mathbf{I} - \hat{\mathbf{d}}_i)\mathbf{f}_i + \mathbf{e}_i \\ \implies \mathbf{x}_i &= \left[\mathbf{I} - (\mathbf{I} - \hat{\mathbf{d}}_i)\mathbf{A}_i\right]^{-1}\left[(\mathbf{I} - \hat{\mathbf{d}}_i)\mathbf{f}_i + \mathbf{e}_i\right] \end{aligned} \tag{3.15}$$

with equation (3.15) allowing us to solve directly for total production, $\mathbf{x}_i$, given a matrix of technical coefficients, and vectors of import ratios, final demands and export demands.

The technical coefficients are known, calculated as in section 3.3.1, the import ratios are known, calculated as in section 3.3.2, and final demands are exogenous. But the vector of export demands is yet to be specified. For now, we will assume it is known, and return to it later once we have derived import demand.

3.4.2 Solving for import demand

Given the total production vector, $\mathbf{x}_i$, which we can now calculate from equation (3.15), and an (assumed to be known) export vector, $\mathbf{e}_i$, we can retrieve the vector of import requirements, $\mathbf{m}_i$, using the import ratios which we outlined in section 3.3.2.

Rewriting in vector notation the definition of the import ratios given in equation (3.4), and using the fact that $\hat{\mathbf{x}}\mathbf{y}$ performs an element-wise multiplication, we have:

$$\mathbf{m}_i = \hat{\mathbf{d}}_i (\mathbf{x}_i - \mathbf{e}_i + \mathbf{m}_i) \tag{3.16}$$

$$\implies (\mathbf{I} - \hat{\mathbf{d}}_i)\mathbf{m}_i = \hat{\mathbf{d}}_i (\mathbf{x}_i - \mathbf{e}_i) \tag{3.17}$$

$$\implies \mathbf{m}_i = (\mathbf{I} - \hat{\mathbf{d}}_i)^{-1}\hat{\mathbf{d}}_i (\mathbf{x}_i - \mathbf{e}_i) \tag{3.18}$$

where each $\mathbf{m}_i$ is an S-vector, with S being the number of sectors.

Equation (3.18) is not defined when a country has zero domestic production for a given sector, leading to a division by zero (or, in matrix terms, to $(\mathbf{I} - \hat{\mathbf{d}}_i)$ having a zero row and hence no inverse). In these cases, the import ratio is set to zero for the purposes of completing the calculation, and the import demand is instead calculated using the expression for total demand, the right-hand side of equation (3.8).

3.4.3 Solving for export demand

In section 3.4.1 and section 3.4.2 we have assumed that the export vector is given. We will now proceed assuming instead that the *import* vector is given and will see subsequently how these two sets of vectors can be solved together using an iteration.

Given a complete set of import demands for each country, given by equation (3.18), we can use the import propensities defined in section 3.3.3 to calculate the export demand this puts upon each country using:

$$e_{is} = \sum_{j=1}^{C} p_{ijs} m_{js} \tag{3.19}$$

We can also solve for all e efficiently using matrix algebra. We start by gathering the import propensities, calculated from data by equation (3.6), into one matrix for each sector, as follows:

$$\mathbf{P}_s = \begin{bmatrix} p_{11s} & \cdots & p_{1Cs} \\ \vdots & \ddots & \vdots \\ p_{C1s} & \cdots & p_{CCs} \end{bmatrix} \tag{3.20}$$

with the importing country indexing the columns and the exporting country indexing the rows. We then further gather these matrices into a block diagonal structure, containing matrices for each sector:

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \mathbf{P}_S \end{bmatrix} \tag{3.21}$$

which has dimensions $SC \times SC$ where, as before, S is the number of sectors and C is the number of countries. The columns of this matrix are a stack of importing countries and sectors, and the rows are a stack of exporters and sectors.

We now arrange the import vectors, given by equation (3.18), in a similar manner, first defining the import demands of a particular sector:

$$\mathbf{m}_s = \begin{bmatrix} \mathbf{m}_{1s} \\ \vdots \\ \mathbf{m}_{Cs} \end{bmatrix} \tag{3.22}$$

then stacking these into a single vector:

$$\mathbf{m} = \begin{bmatrix} \mathbf{m}_1 \\ \vdots \\ \mathbf{m}_S \end{bmatrix} \tag{3.23}$$

The export demands are then calculated as:

$$\mathbf{e} = \mathbf{P}\mathbf{m} \tag{3.24}$$

where the elements of **e** are ordered in the same way as **m**, as a stack of C-vectors, each C-vector representing the export demand of a particular sector.

It should be noted at this point, that the vectors **e** and **m** are not ordered in the same way as in equation (3.18). This implies that care must be taken when calculating the export demand to first reorder the vector of import demands, in order to ensure that we are indeed performing the calculation we expect.

3.4.4 Iteration towards equilibrium

In equation (3.18) and equation (3.24) we have seen how imports and exports depend on one another. This is qualitatively justified by the argument that increased import demand coming from abroad generates a commensurately increased export demand for a particular exporter. This additional export demand generates an increase in the quantity of intermediate input demanded within the exporter which, in turn, requires an increased import demand.

In a standard MRIO model, this cyclical relationship is dealt with using a single matrix inversion. Indeed, the same argument is applied within countries, as increased intermediate demand *for* the goods of one sector leads to an increased demand for intermediates *by* that sector, which justifies the matrix inversion approach for solving a single country's input-output model.

At a global scale, the matrix inversion approach taken by MRIO is undeniably computationally convenient. But given the computing power now available to researchers, this is no longer a sufficient justification for its use, if it is accompanied by assumptions which would not otherwise be made.

The model in the form we present here relaxes none of the MRIO assumptions, but the separation of international trade from individual domestic economies offers the flexibility to do so in future work, simply by making the desired alterations to the global-level model. Such alterations would be more straightforward to perform and more intuitively comprehensible in this model than in an MRIO table, where changing any individual process would require consideration of the entire system and would likely necessitate a completely new approach to finding solutions. Recent work on complexity in economics such as Beinhocker (2006) and Ramalingam et al. (2009) supports the thesis that many of the interesting phenomena in macroeconomics happen when systems are out of equilibrium. With this in mind, we solve the model iteratively, rather than through a single, large matrix inversion.

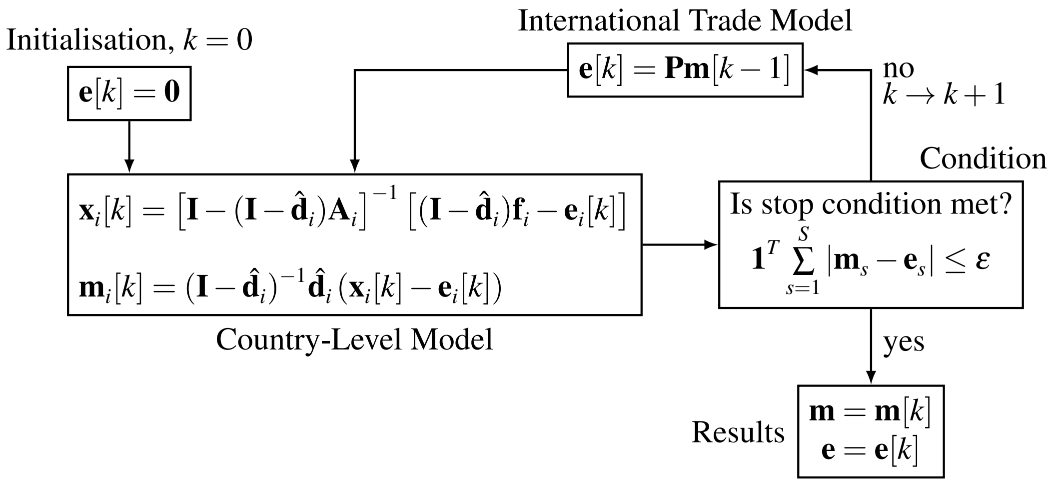


Figure 3.2: The model algorithm: total production, imports and exports are calculated for a given set of fixed coefficients. Note that square brackets have been introduced to designate the iteration variable k, which tracks the quantities that are iteratively recalculated as the algorithm runs. The operator $\mathbf{1}^T$ represents an element-wise sum.

The iterative procedure is outlined in figure 3.2. Firstly, exports in all countries are set to zero. Each country model is solved using these zero exports and a set of import demands is calculated using equation (3.15) and equation (3.18). At this point, a stop condition is applied: does global import demand match global export demand for each sector to within an acceptable tolerance? Clearly, in the first iteration, the answer to this question will be no (since exports have been set to zero).

After the check condition the export demands are recalculated using equation (3.24) and we again calculate import demands and check the stop condition. Notice the order of the operations in the stop condition: each sector's imports and exports are compared in absolute value terms, and the total absolute difference is calculated across all sectors. This ensures that *each sector* is balanced globally between supply and demand. The elements of this resulting C-vector are then summed to get a total global discrepancy between imports and exports across all sectors. This final element-wise sum is represented in figure 3.2 by pre-multiplication by a row vector of ones, $\mathbf{1}^T$.

This process is repeated until the stop condition is true. At this point the model is deemed to have converged and the final import and export vectors are stored. With the 40 countries and 35 sectors of WIOD, this process converges in 10 iterations and takes just under a second to complete on a current desktop computer.

We have thus seen how the heart of the model involves generating a set of imports and exports for each sector in each country from four sets of coefficients:

– final demands,

– technical coefficients,
– import ratios,
– import propensities.

The first of these is given from data (or set exogenously) and the latter three calculated from data. Beyond imports and exports, the model can be used to calculate other metrics which will be useful in some of the analysis which follows.

3.4.5 Model metrics

The two metrics we will use most frequently in the subsequent analysis are "total value added" (TVA) and "balance of trade" (BoT). TVA is a form of gross domestic product (GDP) which is a measure of everything produced in an economy minus that which was required to produce it. This is precisely the "production" estimate method used by the UK's Office of National Statistics in calculating GDP (ONS 2013).

Using the matrix definitions from equation (3.14), we begin by defining a matrix of domestic intermediate flows:

$$\mathbf{Z}_i = \mathbf{A}_i\mathbf{x}_i \tag{3.25}$$

from which we can get a vector of total intermediate usage per sector by summing over the columns:

$$\mathbf{z}_i = \mathbf{1}^T\mathbf{Z}_i \tag{3.26}$$

where, as described in section 3.4.4, the symbol $\mathbf{1}^T$ is a row vector of ones, premultiplication by which is equivalent to summing over columns of a matrix. Finally, by subtracting this value from total production, what remains is a vector of value added per sector:

$$\mathbf{VA}_i = \mathbf{x}_i - \mathbf{z}_i \tag{3.27}$$

By summing this vector across its elements, we arrive at TVA_i, the total value added for country i.

Balance of trade (BoT) is a subset of the more widely-known balance of payments which only measures the difference between total imports and total exports, also called "net exports". Thus when BoT is positive, a country exports more than it imports. Balance of payments includes unilateral transfers such as aid, as well as investment flows. In the current context, we are able to measure BoT per sector in each country. It is measured very simply as:

$$\mathbf{BoT}_i = \mathbf{e}_i - \mathbf{m}_i \tag{3.28}$$

and we can arrive at the total balance of trade across all sectors, BoT_i, by again summing across the elements of the per-sector vector.

3.5 Derivation of country-sector/country-sector flows

In making the assumptions we made throughout section 3.3, we have diverted from the MRIO table which offers a view of trade flows from and to particular sectors in particular countries. For example, an MRIO table might contain an entry recording the flow of German steel to the Chinese vehicles sector. We might term such flows as these as country-sector/country-sector flows, or *CSCS flows*. It will often be useful to have a picture of CSCS flows and, with a simple additional assumption, we can retrieve these from the model structure and return to the MRIO representation from which we diverged in the sections above.

In section 3.3.2 we described an assumption whereby imported goods entered into a warehouse along with all domestically produced goods at which point they became indistinguishable. With the small additional assumption that sectors draw from this warehouse randomly (or, alternatively, that the indistinguishable products are perfectly mixed within the warehouse), it can be argued that the fraction of goods coming from each country of origin will be the same in each sample. We can thus determine the (average) make-up of product origins in each sector's intermediate demand, which is precisely what an MRIO table attempts to do.

The CSCS flow from sector r in country i to sector s in country j can be calculated using the following logic: we know from equation (3.7) that total intermediate demand for good r in the production of s in country j is given by $a_{jrs}x_{js}$. By the definition of the import ratios, a fraction d_{jr} must be imported from elsewhere. And finally, by the definition of the import propensities, and with the assumption made about random sampling from the warehouse, a fraction p_{ijr} of this will come from country i. Thus the CSCS flow is given by:

$$y_{ijrs} = p_{ijr}d_{jr}a_{jrs}x_{js} \tag{3.29}$$

By an identical argument to that used to derive equation (3.29), we can define the part of final demand for sector s in country j which comes from country i as

$$f_{ijs} = p_{ijs}d_{js}f_{js} \tag{3.30}$$

These CSCS flows will be useful later on when we come to join the model to other social science models such as of migration and of international development and also as a method of assessing what impact a given change to the model's coefficients has.

3.5.1 A comparison with a WIOD MRIO

Using equations (3.29) and (3.30) we can reconstruct an MRIO by assembling y_{ijrs} and f_{ijs} flows into an appropriate matrix structure. By doing this we can compare value for value against an MRIO to encourage ourselves that the process of converting a static MRIO into a model which we have described so far in this chapter, does not have too severe effect on the underlying structure.

Clearly there is some loss of information here. As discussed above, we are losing fidelity when we convert data points to coefficients for modelling purposes. But our goal in this section is to show that the information lost does not change the resulting MRIO structure in a very severe way and that the two ways of representing the global economy are recognisably representing the same world.

A very simple way to measure the similarity of the two systems is to calculate the (Pearson's) correlation coefficient which is a measure of the covariance of two random variables scaled by the standard deviation of each. A value of 1 indicates that the two variables are perfectly correlated and 0 indicates no correlation. By rearranging our two matrices as ordered vectors (where, clearly, care is taken to ensure the two vectors are ordered identically), we can calculate the correlation coefficient of the WIOD MRIO and the MRIO resulting from the demonstration model. The coefficient turns out to be 0.97.

Another method of comparing the two MRIOs is to make a visual representation of each and compare them by eye. Because of the large number of values involved, this is quite a fruitful way of broadly comparing structural similarity.[9]

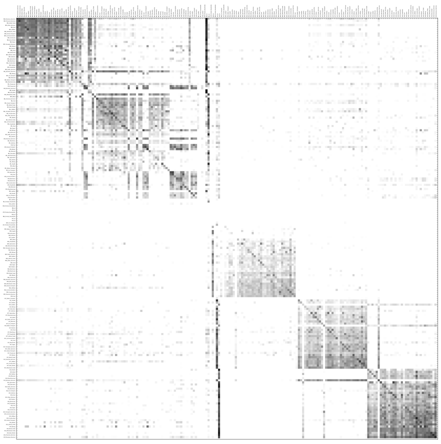
To achieve this, we plot each matrix as a heat map, with higher values being represented by darker colours. Additionally, to help show structure, we first order the rows and columns using a sorting algorithm called 'seriation' which seeks to minimize the gradient at each point of the heat map (i.e. maximise the 'smoothness'), both vertically and horizontally. The algorithm is implemented in R (Hahsler et al. 2015). This is a novel analysis method in the context of global trade modelling, and one which gives a useful way of quickly getting an overview of large systems. This will be a theme throughout the analyses of this thesis.

As the matrices being compared are large, we present three different views of the WIOD MRIO and the model MRIO for comparison. Figure 3.3 shows six heatmaps for visual comparison. The images in the left column correspond to MRIOs directly from WIOD,

[9] There are $(S \times C)((S+1) \times C)$ values which is just over 2 million with the WIOD system of sectors and countries

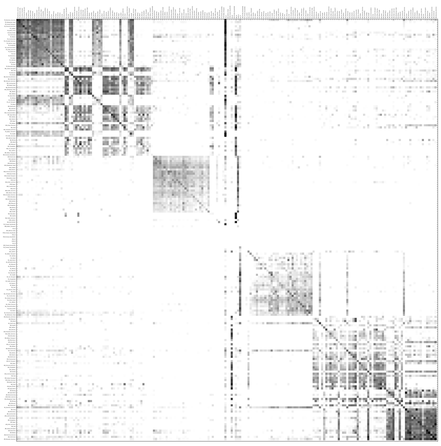
and those on the right are the MRIOs reconstructed from the demonstration model. In the top row, the first 200 rows/columns of each MRIO (where the order is specified by WIOD, country ISO3 codes alphabetically, then sectors according to NACE code, followed by final demand). The middle row shows 200 rows/columns from the central portion of each MRIO ordered as before, and the last row shows a random sample of 200 rows/columns to ensure any similarity is not local to the other two areas chosen.

Although this is a somewhat blunt instrument, visually, we can have good confidence that the overall shape and structure of the MRIO is retained during the modelling process. Indeed there are only subtle clues that the images on the left are not identical to those on the right (for instance, the left-most fifth of the top row, one-fifth down from the top is clearly darker in the WIOD MRIO than in that of the model.) We therefore conclude that the great majority of the structural information from WIOD is retained by this reduced-information model.

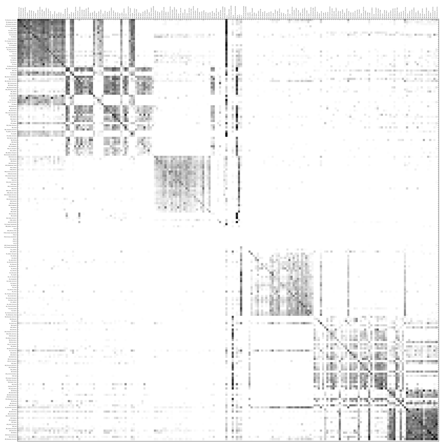


(a) WIOD MRIO (first 200 rows)

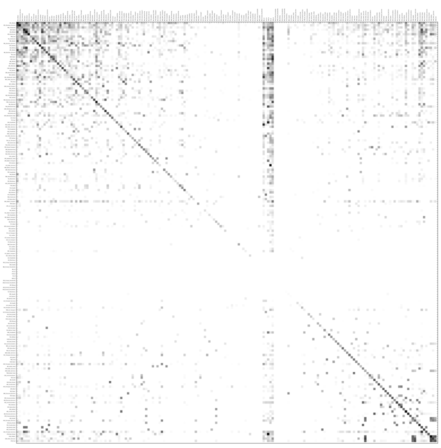
(b) Model MRIO (first 200 rows)



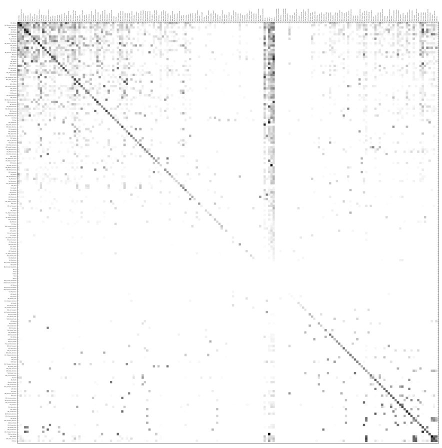
(c) WIOD MRIO (middle 200 rows)



(d) Model MRIO (middle 200 rows)



(e) WIOD MRIO (random 200 rows)



(f) Model MRIO (random 200 rows)

Figure 3.3: A visual comparison of the MRIOs from WIOD data (left column) and MRIOs reconstructed using the demonstration model (right column). Seriation (i.e. row and column ordering) was performed on the first 200 rows (top row), the middle 200 rows (middle row) and a random sample of 200 rows (bottom row).

3.6 The rest of the world

All COMTRADE flows coming from regions, or from countries not in the demonstration model are deemed to have come from what we might generically term "the rest of the world" (RoW). We describe flows either to or from RoW as being "stray".

RoW exists in the model as a country like any other, but has some special considerations since, clearly, it has no associated input-output table.

Any country which is reported in COMTRADE as importing zero of a particular sector, but whose input-output table requires there be some imports of a particular sector available, is deemed to get its product entirely from RoW.

The final demand vector for RoW is calculated at model set up time by summing all stray exports. By indexing variables with ρ to indicate from/to RoW, we can write this as:

$$f_{\rho s} = \sum_{i=1}^{C} \bar{y}_{i\rho s} \tag{3.31}$$

where, as before, a bar indicates a quantity taken from data. Equation (3.31) implies that RoW consumes everything it imports, and that it demands just enough to 'soak up' stray exports from the countries which are in the model.

RoW has neither technical coefficients nor import ratios. Instead it can produce goods "for free", i.e. without inducing any intermediate or import demand. Thus, instead of calculating total production via equation (3.15), total production in RoW is simply set to be equal to export demand. Thus, we can write:

$$\mathbf{x}_{\rho} = \mathbf{e}_{\rho} \tag{3.32}$$

where $\mathbf{e}_{\rho}$ is calculated via the iterative procedure given by figure 3.2 exactly as for any other country (with the caveat that RoW's import demand vector is 'decoupled' from its export vector by the fact that RoW can produce sufficient goods to satisfy any export demand for free.)

This concludes our description of the functioning of the "rest of the world" entity, and with it, our technical description of the model as a whole.

3.7 Conclusion

In this chapter we have described in detail a new model of the global economy, motivated by a discussion of the need to restrict economic assumptions in favour of empirical observation, to make available to modellers from other social science disciplines a set of adjustable coefficients for use in their particular modelling scenarios, and to be as parsimonious as possible in terms of the data requirements at the country level in order to facilitate the addition of countries to the model which publish little or no comparable country-level data.

The model is a 'happy medium' between the complexity of CGE/GTAP and the linear simplicity of MRIO. Much of the structure reflects the earlier work of Project LINK, who wanted similarly to build a demonstration model in order to make broader points about the benefits of modelling globally.

Our model consists of country-level descriptions in the form of input-output tables taken from the World Input-Output Database (WIOD), combined with a description of the global trade network taken from the UN's databases of commodities (COMTRADE) and services (ServiceTrade). We kept to the WIOD categorisation of the economy into 35 sectors and outlined how to translate from the UN data sources to the WIOD sectors.

We described the four sets of coefficients needed to completely describe a country and its place in the global trade network, namely: final demand levels, technical coefficients, import ratios and import propensities. We then went on to show how these four sets can be combined to calculate, via an iterative procedure, a complete set of intermediate flows, imports and exports for each country in the model and for the "rest of the world" entity. The use of iteration to solve the model was justified in terms of the extensibility given to the model by this less mathematically elegant but more flexible approach, with particular emphasis given to the potential in future of introducing Armington imperfect-substitute assumptions, and relaxing the assumptions of linear production functions and fixed linear import propensities.

We then introduced two metrics, total value added and balance of trade which will go on to be at the heart of the analysis which makes up the remainder of this thesis. We showed how these metrics are to be calculated from the four sets of model coefficients. Also useful to the analysis which follows, we showed how our model can be reduced to a traditional multi-region input-output model by calculating country-sector/country-sector flows by making some additional assumptions.

The casual reader is encouraged to skip the remainder of this part, which deals with improvements to the model as outlined here, and move straight to part III where the model is used in a series of analyses, both "within-model" (i.e. studying the properties of the model itself) in chapter 6 and using the model as part of a wider modelling question. Examples of how the four sets of model coefficients can be used in other social science contexts are given in chapter 8 (migration), chapter 9 (international security) and chapter 10 (international development).

The remainder of this section discusses extensions to the model as discussed in this chapter. We first turn to an improved method of estimating services flows, in chapter 4. The addition to the model of the 150 or so countries which do not publish input-output data in WIOD is then assessed in chapter 5.

Chapter 4

Improving the estimation of services flows

4.1 Introduction

The model presented in chapter 3 asks a lot of the data it is provided with. The magnitude of each country-country-sector flow, y_{ijs}, is used to calculate an import propensity, p_{ijs}, which is then fixed and determines, to a large extent, the response of the model to an exogenous change in final demand. Each piece of data is taken as telling us something fundamental about how the global trade network operates, and the relative structure (but not, admittedly, the absolute magnitude) of the network is derived entirely from observed quantities.

In the case of commodity flows (i.e. physically tradeable goods), this commitment to the primacy of observation is a natural way to proceed; due to the need to impose customs taxes/tariffs on physical goods entering or leaving a country, the trade in commodities is extremely well recorded.

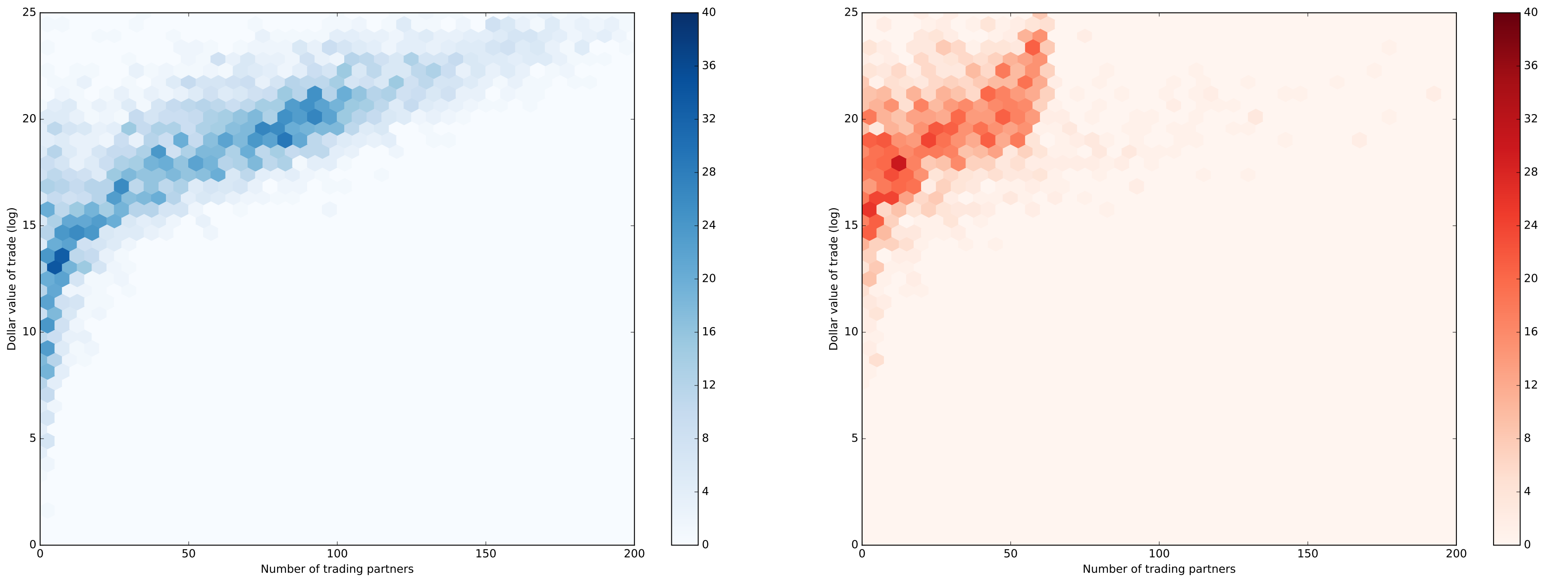
The same cannot be said of trade in services which, since it is not done via border controls, is far less accurately recorded and often not recorded at all (Dietzenbacher et al. 2013, p. 86). Indeed, one could argue that it is not clear what trade in services even *is*, but answering that tricky economics question is outside the scope of this thesis (Francois and Hoekman (2010) include a discussion of some of the issues involved in measuring and defining trade in services).

Instead, we will content ourselves to follow the WIOD in assuming that services can be both imported and exported in a manner identical to that for commodities. We will also assume that the UN ServiceTrade database is the best available reflection of the extent to which this trade occurs globally. Again, we will choose here to take this data at face value, despite the fact that we might have cause to suspect that services trade is incompletely reported, or that there may be other sources of trade data (the OECD and the GTAP projects are conspicuous examples).

In many cases in the services trade data, *total* imports or exports for a particular country-sector are reported, rather than the country-country-sector flows which our model requires for its initialisation. Or, where point-to-point flows *are* reported, the reported trade partner is sometimes a region, rather than a country. Figure 4.1 shows the difference between the commodities and services data sets. Each graph shows a density plot for one entire dataset for 2010. This is another example of how to visualise large datasets and models, which is a key theme of this thesis. As far as we know, this is first use of density plots of this kind in trade modelling literature.

The flows are divided into "products", namely into WIOD sectors for the commodities flows, and individual UN service codes for the services data. Since the services are far more broadly categorised than the commodities, these two categorisations are loosely equivalent in terms of the numbers of categories they produce. The total reported dollar trade for each product is calculated, and graphed against the number of individual trade records making up that product. Thus products where many of the records are the kind of country-sector trade totals outlined above will tend to appear towards the left or top-left of the graph, and products for which many individual country-country flows are reported will tend to appear towards the right.

According to this broad rule of thumb, the services do indeed seem more to consist of trade totals than do the commodities. In this chapter we will therefore assume that only trade totals (i.e. total exports and/or total imports) are known. We will be concerned with methods for estimating country-country-sector flows from trade totals, which is a fairly well understood problem. Despite this, as far as we know, none of these approaches has ever been applied to the UN ServiceTrade database and this is therefore a novel contribution of this thesis.



(a) Commodities

(b) Services

Figure 4.1: Density plots showing the total (logged) dollar value of trade for each category of either good or service, against the number of bilateral importer/exporter combinations which have non-zero trade in each category.

4.2 Estimation via iterative proportional fitting

As we have seen, this chapter concerns itself with the estimation of point-to-point flows when only flow totals, either in- or out-flows, are known. Iterative proportional fitting can be seen as a true least-assumptions approach to this problem in that it treats the point-to-point flows purely probabilistically when the initial matrix has no information (more on this below). It is identical in this situation to solution via entropy maximisation and the 'null model' in community detection as we will see later. One caveat is that both margins must sum to the same total if the algorithm is to converge.

4.2.1 The method

Iterative proportional fitting starts with an initial matrix and a set of known row and column totals, which must sum to the same value. The initial matrix may be a set of measured flows which are considered to contain relevant information on the relative sizes despite the fact that they may not sum to the correct margin totals. Alternatively, in the complete absence of point-to-point information, the process starts with a matrix of ones.

In either case the method is the same and consists of an algorithm which iterates between row operations and column operations. In each row operation, all rows are made to sum to the required margins by simply dividing by the current row sum and multiplying by the margin in question. Thus,

$$f_{ij}^{(1)} = \frac{f_{ij}^{(0)} O_i}{\sum_k f_{ik}^{(0)}} \tag{4.1}$$

where $f_{ij}^{(n)}$ is the matrix entry at ij after the nth iteration, and O_i is the ith row sum. This is repeated for the columns. This row-column iteration is then repeated until the row and column sums are both simultaneously within some arbitrarily small distance from the desired values according to some measure. For example, we might measure the sum of squared deviations.

There are only two further points to make about the method. The first is that since only multiplications and divisions are involved, any zeros in the initial matrix will remain zero throughout the procedure. This will be useful in cases of estimating trade flows if countries/regions etc. are disallowed from trading with themselves, in which case the diagonal elements of the initial matrix can be set to zero.

Secondly, in the case where the initial matrix contains some information (i.e. where the non-zero initial values are not all the same), it is generally the case that, all else being equal, larger initial values tend to result in larger final values. Therefore IPF can, in some very weak sense, be seen as rank-preserving. To demonstrate this empirically, we take an initial matrix of random values between zero and one. We then run IPF against randomly chosen row and column totals. The only constraint on the row and column totals is that they were both forced to sum to the total sum of the randomly selected initial matrix values.

Figure 4.2 shows the relationship between initial value and final value in this experiment with 10,000 runs and a random initial matrix of 5×5. Each matrix value had its initial value tracked, and paired with its value after iteration had completed. These initial/final pairs can then be scatter-plotted with the initial value on the x-axis and the final value on the y-axis. For visual clarity, figure 4.2 is not a scatter plot, but rather a density plot, where the xy space is binned into hexagonal regions. The darker hexagons represent more data points falling into that area of the space.

To show that larger values generally end up larger after iteration, y should be correlated with x and we should expect to see an upward-sloping distribution with more values at the bottom-left than at the top left, and more values at the top-right than at the bottom-right. This particular run has a Pearson correlation coefficient of 0.3 (with a very small p value, since n is of the order 10^5) between the initial and final values, demonstrating that larger initial values tend to lead to larger final values.

In fact, Mosteller (1968) showed something rather more subtle than this: that IPFP retains the "interaction structure" of the original matrix, defined as a certain crossproduct ratio. For our purposes, our simpler definition will suffice and we will assume that IPF broadly maintains initial matrix-value rankings.

4.2.2 With all initial values equal

If there is no initial information on which to base an estimate of point-to-point flows, the initial matrix is usually set to have all values equal to one. To see how this affects the result, we solve equation (4.1) when all $f_{ij} = 1$. The equation for the first iteration is then:

$$f_{ij}^{(1)} = \frac{O_i}{n} \tag{4.2}$$

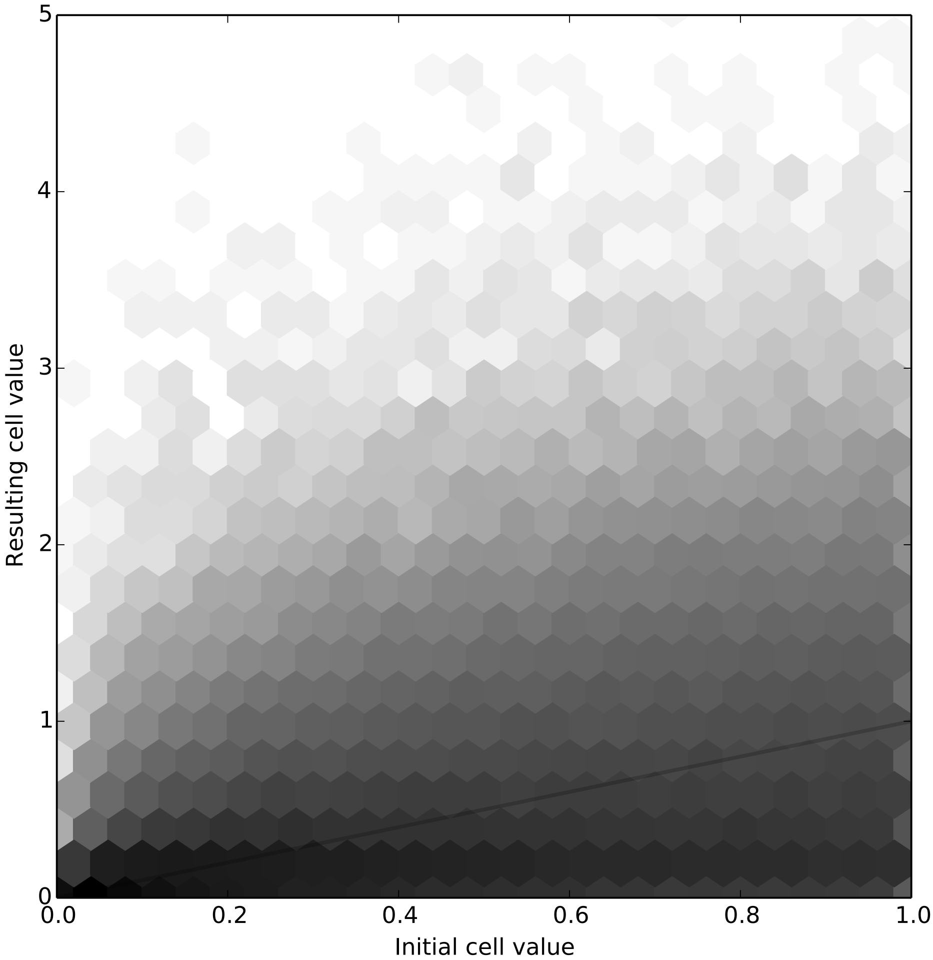


Figure 4.2: Density plot showing the results of multiple iterative proportional fitting runs. During each run, the initial value of every matrix element was tracked and paired with its value after the iteration completed. This graph shows these pairs of initial (x-axis) and final values (y-axis), having first binned the initial/final space into hexagonal shaped regions for visual clarity. Darker hexagons represent more matrix elements occupying that space. The initial matrix was uniformly distributed between zero and one with shape 5×5. 10,000 runs were performed. The straight line shows $y = x$.

where n is the dimension of the square matrix. We then run the column step of the iteration:

$$f_{ij}^{(2)} = \frac{f_{ij}^{(1)} D_j}{\sum_k f_{kj}^{(1)}} = \frac{O_i D_j}{\sum_k O_k} \tag{4.3}$$

Since this method has a closed form solution[1], the algorithm always completes in a single row/column-operation. This result has some interesting equivalences in seemingly unrelated fields.

4.2.3 Relationship to graph theory

Since the row and column margins are defined as summing to the same value, we can represent both $\sum O_i$ and $\sum D_j$ as a single value a. We can then rewrite equation (4.3) as

$$f_{ij}^{(2)} = \sum_k D_j \frac{O_i D_j}{\sum_k O_k \sum_k D_j} = a \frac{O_i D_j}{\sum_k O_k \sum_k D_j} = a A_i B_j O_i D_j \tag{4.4}$$

where we define $A_i = \frac{1}{\sum O_i}$ and $B_j = \frac{1}{\sum B_j}$.

From this last representation, it is clear that when the initial matrix contains no information, IPF is identical to entropy maximisation (up to the multiplicative factor a) as presented in, for example, Wilson (1969). The term entropy maximisation is equivalent to the more common 'maximum likelihood.' Given the margin totals, IPF gives us the maximum likelihood flow matrix, given no additional information.

Also interesting to note is the relationship between IPF and some of the tools of graph theory which is used to analyse networks of nodes and edges. This work will inspire the analyses of part III where clustering and community detection algorithms are used a number of times, in ways which are novel to the trade modelling literature. See, for example, pp. 149, 185 and 283.

If we define the proportional limits as being:

$$p(M_i) = \frac{M_i}{a} \tag{4.5}$$

[1] Provided, as mentioned above, that the margins sum to the same number. If not, either the rows or the columns will sum to the correct margin depending on which is done last. The other margin will sum to an arbitrary number.

where M_i is a margin total, either O or D, then equation (4.3) can also be written as

$$f_{ij}^{(2)} = ap(O_i)p(D_j) \tag{4.6}$$

which is the same as the second choice of null model proposed by Reichardt and Bornholdt (2006b, p.3) for assessing the fitness of a particular community grouping in network analysis. In their context, O_i is a node's *in-degree*, the number of directed edges pointing in to the node, and D_j is the *out-degree*. Crucially, it is this choice of null model which makes the Reichardt and Bornholdt (2006b) method identical to that of Newman and Girvan (2004a) who first defined the critical concept of *modularity* which is at the heart of much subsequent work in that literature. Rosvall and Bergstrom (2008, p.1122) also refer to this choice of null model and describe it as "the most general form" of the modularity maximisation approach.

4.2.4 Estimation with some known flows

When some flows are known, but not a complete matrix, the simplest approach recommended by Lahr and Mesnard (2004, p.125) is simply to subtract the known flows from the relevant margin totals and set the entries in the initial matrix to zero.

By doing this, the known flows remain zero, and the remaining flows are proportionally fit to the margins *without* the known flows such that the known flows can simply be added back in at the end.

Notice that this process makes explicit the fact that a known flow has no influence over any other point-to-point flow other than through its effect on the relevant margins. A more sophisticated approach to dealing with known flows would require a model for how one point-to-point flow affects other such flows. Such a model is beyond the scope of this thesis.

4.2.5 Drawbacks to estimating services flows with IPF

As discussed above, IPF is very much a 'no-information' solution to the problem of estimating point-to-point flows from flow totals. Services flows estimated in this way will contain none of the trade patterns we might expect to see in a global commodities trade network. For example, if the Netherlands is a large exporter of a particular service then this will be reflected equally in the import patterns of both the UK and China. But this disallows for the possibility that countries within the EU have greater-than-proportional trade when compared with that which occurs across EU borders, which is highly probable. It also ignores the well-established fact that, as with gravity modelling of commodities, distance between trading partners is also a significant determinant of services trade volumes (Kimura and Lee 2006).

The problem with using IPF to estimate services flows is that we might be discarding important information about patterns of global trade contained in other data sources. The next section seeks to address this problem by making some assumptions about similarities between the services trade network and that of trade in commodities.

4.3 Estimation via the gravity model

In the study of global trade, we are far from being in the no-information situation which we assumed in our discussion of IPF in the previous section. As mentioned previously, trade in physical commodities is extremely well documented. In this section we will use the information contained in the commodity flows to make better estimates than we were able to using IPF. We will begin with the simplest possible exposition, and add further complexity until we are satisfied that we have used as much of the information as we have available as feasible.

4.3.1 The method

In section 4.2.5 we discussed the confounding effect which political trade relationships will tend to have when estimating flows purely proportionally: countries within the trade relationship will trade more with one anther than they "should" if trade were determined proportionally. Similarly, the most commonly cited factor which distorts trade from pure proportionality is geographical distance. For example Disdier and Head (2008) perform an excellent meta-analysis of the literature on how distance affects trade, and the literature review in chapter 2 covers the topic in detail.

When combined with a bi-proportionality related to the GDP of the importer and exporter country, as does McCallum (1995), this model is called the gravity model and remains a standard tool in trade analysis. For example, as recently as 2013 it was still being used in its very simplest form by the World Bank to estimate trade costs Novy (2013). The model is usually presented as:

$$y_{ij} = \frac{Y_i^{\alpha} Y_j^{\beta}}{d_{ij}^{\gamma}} \tag{4.7}$$

where Y_i is the GDP of country i and d_{ij} is the geographical distance between the two countries. The parameters α, β and γ are to be specified from data.

For the present purposes we will use minimum distances (such that neighbouring countries have $d_{ij} = 0$) from the CShapes database (Weidmann, Kuse, and Gleditsch 2010).[2] Zeros in the data present a problem for multiplicative formulations such as this one, since division by zero is not defined, and multiplication by zero always results in zero

[2] For an exhaustive discussion on other potential measures of distance, such as straight line between capital cities, see Frankel (1997). The crucial finding of this work is that the many possible distance measures tend to give broadly similar results.

overall. But here we will simply replace d_{ij} with $e^{d_{ij}}$ throughout the remainder of this treatment. When using centroid distance, this replacement will not, in general, be necessary.[3]

The parameters α, β and γ are to be estimated from some other known dataset, for example, data from a previous year. Due to the multiplicative nature of this equation, the estimation is easy to carry out by ordinary least squares regression by taking logs.

Usually the formulation in equation (4.7) is extended to include other variables which affect levels of trade. Written more generally, if there are K variables thought to affect trade, we have:

$$y_{ijs} = \prod_{k=1}^{K} x_{kijs}^{\beta_k} \tag{4.8}$$

where x_{kijs} is the kth variable relating trade from i to j in sector s. There are then K parameters to be estimated. Clearly, denominator parameters such as distance in equation (4.7) are represented by having a negative parameter. Finally, variables which do not relate to one or other of i or j are fixed to be the same for all values. For example, in the case of exporter GDP:

$$x_{kij} = x_{ki} = Y_i \quad \forall\, k, j \tag{4.9}$$

Table 4.1 shows the results of using equation (4.8) with a variety of specifications[4]. The results of the formulation in equation (4.7) are shown in column 1.

The minimum distance term has the expected negative sign (the regression was run multiplicatively, i.e. with d_{ij} in the numerator), and is significant at 1%. The point estimate of -0.35 is somewhat smaller than those obtained by Baldwin and Taglioni (2006), McCallum (1995) and Anderson and Wincoop (2003) who get around -0.8 using a different data set and definition of distance. This discrepancy is not explained by using exponentiated distance: the point estimate is even closer to zero when using log distance (results not shown).

[3] Helpman, Melitz, and Rubinstein (2008) have an elegant, but theoretically complex, method of dealing with this problem in the context of zero trade flows. They explicitly model how countries choose to trade with one another in the first place. See the literature review of chapter 2 for more complete details.

[4] Note a few country/sector combinations have, presumably erroneously, all zero technical coefficients. This would imply that they are able to produce output with no intermediate input. Since this clearly cannot be the case, these country-sectors have been excluded from the analysis. They are CYP Fuel, LUX Fuel and Leather, LVA Fuel and SWE Leather.

Table 4.1: Linear regressions to estimate the parameters of a gravity-type model, to determine the flow from country i to country j of sector s. Six different specifications are shown. Standard errors are in parentheses. Statistical significance is indicated by asterisks.

	Dependent variable:					
	$\log(y_{ijs})$					
	(1)	(2)	(3)	(4)	(5)	(6)
Dummies	none	none	none	none	c	c and s
mindist	-0.345^{***}	-0.346^{***}	-0.344^{***}	-0.342^{***}	-0.348^{***}	-0.353^{***}
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
$\log(x_i)$	1.049^{***}	0.564^{***}	0.576^{***}	1.031^{***}		
	(0.011)	(0.015)	(0.015)	(0.011)		
$\log(x_j)$	0.890^{***}	0.893^{***}	0.892^{***}	0.889^{***}	0.896^{***}	0.904^{***}
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)
$\log(f_{is})$		0.485^{***}	0.463^{***}		0.511^{***}	0.368^{***}
		(0.011)	(0.011)		(0.012)	(0.022)
$\log(v_{is})$			-0.486^{***}	-0.870^{***}	-0.160^{***}	-0.267^{***}
			(0.049)	(0.049)	(0.051)	(0.066)
Constant	-22.671^{***}	-19.894^{***}	-20.401^{***}	-23.353^{***}	-12.113^{***}	-11.792^{***}
	(0.206)	(0.208)	(0.214)	(0.208)	(0.190)	(0.269)
Observations	23,913	23,913	23,913	23,913	23,913	23,913
R^2	0.310	0.329	0.359	0.347	0.382	0.650
Note:						*p<0.1; **p<0.05; ***p<0.01

Both the total production figures, x_i and x_j, have the expected positive sign and are highly significant. They are both much closer to the McCallum (1995) estimates of around 1.1 for x_i and 1.0 for x_j.

In columns (2) and (3) of table 4.1, additional independent variables are added which are not available in ordinary trade data sets, since they come from the WIOD input-output tables at the heart of our model of trade. The first is the final (domestic) demand in for sector s in the exporter country. Krugman (1980b, p.956) provides theoretical reasons why a country with a large domestic market for a product might be a larger exporter of that product. Thus we would expect a positive coefficient on f_{is}. The coefficient is indeed positive and highly significant. Interestingly, the inclusion of domestic demand greatly reduces the parameter estimate on total production in the exporter country, x_i, suggesting that some of the production effect in models such as equation (4.7) might really be due to the fact that domestic demand for that product is high (which, of course, will have a direct effect on total production.) This is an interesting and, we believe, novel result.

Another determinant of the attractiveness of an exporter might be its ability to produce good s cheaply. A proxy for this, added in column (3), is the value added per unit output, defined from the technical coefficients for the relevant sector as:

$$v_{is} = 1 - \sum_r a_{rs} \tag{4.10}$$

This makes intuitive economic sense since, under perfect competition assumptions, price is equal to the marginal cost of production, defined as the price to create a single unit of a good at current production levels. The technical coefficients represent the intermediate costs of producing a unit of good, although they do not include labour costs. It is in this sense that they are *a proxy for* production costs rather than being the production costs themselves.

All things being equal, we would expect countries with a lower production cost to export more, since they are more efficient producers. The negative sign on v_{is} in table 4.1 is therefore surprising. Various different specifications (not shown) have convinced us that this negative coefficient on v_{is} is very robust, showing that this is not merely a misspecification problem. We will investigate this further in section 4.3.2.

Columns (4) and (5) of table 4.1 show the effect of adding country (c) and sector (s) fixed effects. With regards to country fixed effects, Anderson (2010, p.27) states that excluding them leads to a host of common econometric problems, such as misspecification and omitted variable bias. Of course, with country fixed effects included, x_i must be excluded since it is also fixed per country. The general result though is that the results discussed above are all robust to the addition of sector and country fixed effects: no point estimate is changed hugely, other than the coefficient on v_{is} which is greatly reduced (although not in order-of-magnitude terms), and none of the variables loses its significance level.

4.3.2 Splitting up value added

We now investigate further the result that v_{is} has a robustly negative effect on an exporter's trade attractiveness. We begin by splitting v_{is} into its constituent r parts $1 - a_{rs}$ in equation (4.10). Table 4.2 shows the result of including each of the "from sector" technical coefficients (restricted to the commodity sectors). The regression run was identical to that in column (5) of table 4.1, sector dummies being inappropriate here for obvious reasons. The independent variable corresponding to each row of the table can be thought of as a measure of how much of each sector was used in total per-unit production in the exporting country.

We might expect that lower technical coefficients lead to more trade, since this implies a more efficient production technology[5]. The picture from the per-sector analysis is rather mixed. Of the sectors with significant ($p < 0.01$) parameters, leather, fuel, metals, minerals, paper and textiles all have the expected negative sign: a production technology which is efficient in these sectors encourages export. Particularly important here are metals, minerals and paper which all have comparatively large point estimates. It is perhaps encouraging that the coefficient on fuel is significant and negative. This suggests that fuel-efficiency is indeed a factor when importers are choosing whom to trade with.

But there also sectors with a positive and significant ($p < 0.01$) parameter. These are plastics, machinery and chemicals, and the point estimates on these sectors are also comparatively large. Perhaps we might conclude that the difference comes from the extent to which these sectors require inputs themselves. Metals, minerals and paper are all close to being raw materials, requiring little input in their manufacture beyond raw materials and energy. We will refer to these as primary products. But plastics, machinery and chemicals are secondary products, requiring a wider range of inputs to produce. It seems that using comparatively more primary product in a production process is a *discouragement* to export, but that using comparatively more of a secondary

[5] It does in intermediate terms at least. Labour is not included in this measure and is implicitly assumed to be equal across all sectors and countries; a highly questionable assumption. Further research here might illuminate whether including labour costs, which are available from WIOD, explains the mysterious negative sign on v_{is}.

product is an *encouragement* to export. Perhaps what we are observing here is a preference for refined goods over unrefined, certainly in the dollar value terms which all these flows are measured in. An enhanced version of this analysis involving the addition of a labour market model would be a good first step in moving the model towards the complexity of a full CGE model.

The regression coefficients given in table 4.1 can be used as the basis for a gravity model of services trade specified as in equation (4.8). We can use this model to estimate point-to-point services flows just as we did by using IPF in section 4.2. The obvious next step is therefore to compare these two methods of estimating services flows, and it is to this task which we now turn.

Table 4.2: Regression similar in structure to table 4.1 but with v_{is} divided into its constituent sectors.

	Dependent variable:	
	$\log(y_{ijs})$	
	(1)	(2)
log(food)	0.088***	
	(0.021)	
log(leather)	−0.046***	
	(0.015)	
log(fuel)	−0.081***	
	(0.026)	
log(plastics)	0.104***	
	(0.024)	
log(metals)	−0.181***	
	(0.027)	
log(machinery)	0.127***	
	(0.025)	
log(vehicles)	−0.043*	
	(0.024)	
log(agriculture)	−0.031***	
	(0.011)	
log(wood)	−0.029	
	(0.018)	
log(minerals)	−0.107***	
	(0.020)	
log(manufacturing)	0.027	
	(0.021)	
log(paper)	−0.236***	
	(0.028)	
log(electricals)	0.084***	
	(0.027)	
log(textiles)	−0.059***	
	(0.021)	
log(mining)	0.015	
	(0.015)	
log(chemicals)	0.124***	
	(0.030)	
log(total)		−0.012
		(0.061)
$\log(x_j)$	0.893***	0.893***
	(0.009)	(0.009)
$\log(f_{is})$	0.328***	0.350***
	(0.023)	(0.022)
Observations	23,019	23,019
R^2	0.617	0.612
Note:	*p<0.1; **p<0.05; ***p<0.01	

4.4 IPF vs. gravity: a comparison of the methods

In section 4.2 we were introduced to the iterative proportional fitting (IPF) method of estimating services flows, and section 4.3 demonstrated how commodity flows could be used to calibrate a gravity model for the estimation of services flows.

We will compare estimates of point-to-point services flows using the both these methods, assuming that the only information available is the import and export totals (row/-column sums), ignoring any point-to-point information we may have. We can then calculate a mean squared error from those point-to-point flows which *are* contained in the UN ServiceTrade database, and use these to compare the methods.

We will restrict ourselves to concentrating on the 40 countries of the WIOD to make the analysis manageable in terms of number of countries.

4.4.1 Unbalanced row and column margins

Since any balancing procedure only makes sense if the row and column margins sum to the same number, and since this would not be expected, in general, to be the case in a set of trade data from disparate reporters, throughout this analysis we will adopt the following procedure to ensure the row and column margins balance.

- Each set of margins will be given an additional element, labelled "RoW" (for "rest of world").
- If the row margins sum to more than the column margins, the RoW element of the row margins will be set to zero, and that of the column margins set to (the absolute size of) the difference between the two sums.
- If the column margins sum to more than the row margins, the opposite procedure will be applied.

4.4.2 Iterative proportional fitting

Since we do not have a complete set of prior information about the point-to-point flows in the services sectors, we will start the procedure with a matrix of ones. In this case, and given the adjustment described in section 4.4.1 to ensure the margins sum to the same number, there is a closed form solution as per equation (4.3).

The column labelled "ipf" of table 4.3 shows the root mean squared error (RMSE) for the iterative proportional fitting method applied to each of the services sectors. The columns of the table are ordered by total RMSE left-to-right from lowest to highest. Iterative proportional fitting is in the middle column.

Table 4.3: The root mean squared error (RMSE) for each services sector when point-to-point flows are estimated using the method associated with each column. RMSE is calculated as $\sqrt{\sum_i (\bar{x}_i - x_i)^2 / n}$. The specification number refers to the column of table 4.1 used as the gravity model coefficients.

	gipf4	gipf3	gipf1	gipf2	gipf5	ipf	gravity4	gravity1	gravity3	gravity5	gravity2
Air Transport	372	373	374	374	375	330	555	535	538	531	532
Business Services	787	787	787	787	787	1206	2201	2127	2091	2090	2021
Communications	130	130	130	130	130	156	269	348	423	558	579
Education	40	40	40	40	40	70	185	308	554	1012	880
Financial Services	294	295	295	295	296	323	840	838	900	1011	1071
Health	9	9	9	9	9	13	145	265	824	1233	1222
Hospitality	1113	1111	1111	1110	1109	1634	2625	2570	2406	2393	2326
Inland Transport	218	218	218	218	218	237	436	468	492	507	573
Other Services	988	989	989	990	991	1253	2148	2084	1934	2011	1862
Public Services	94	95	95	95	95	101	549	586	1345	1704	1962
Real Estate	29	29	29	29	29	42	96	207	601	920	991
Retail Trade	644	644	644	644	644	732	1356	1314	1302	1201	1404
Transport Services	250	249	249	249	249	333	516	530	505	524	505
Utilities	70	71	71	71	71	75	217	282	393	460	516
Vehicle Trade	67	68	68	68	68	67	140	250	271	304	398
Water Transport	578	579	579	580	582	413	770	771	810	818	808
Wholesale Trade	64	64	64	64	64	67	142	282	551	830	832
Total RMSE	5748	5751	5751	5753	5755	7051	13191	13765	15941	18108	18482

The sectors vary a lot in terms of their root mean square error, but this is likely due to the fact that the magitudes of the total dollar values of trade of each sector are very different, as shown by table 4.4.

The best-estimated sectors by the RMSE measure are: health (896), real estate (2,148), vehicle trade (2,790), wholesale trade (2,790) and education (1,382). The worst-estimated are: hospitality (2,914), other services (3,092), business services (3,421), retail trade (2,794) and water transport (2,976).[6]

4.4.3 Adding a gravity model

We will now explore each of the gravity-model specifications given in table 4.1, estimated using commodities flows, to see which produces the set of flows which most closely matches those point-to-point flows which *are* in the dataset. This is an *in-sample* comparison. Specification (6) is not included here since the sector dummies in table 4.2 are for commodities sectors and we are now dealing with services sectors.

The results are shown in the columns labelled ‘gravity1’ to ‘gravity5’ in table 4.3. We will start with specification (1), the first and simplest, including only distance, and the total production of the importer and exporter. This is the gravity model in what might be termed its “purest” form. It is the second most effective specification behind only (4) which we will look at in more detail below. Along with all the other specifications, this specification varies greatly in how well it predicts the flows of each sector. The magnitudes of the prediction errors seems broadly to follow those associated with iterative proportional fitting. Figure 4.3 confirms that the estimation errors are indeed very closely related.

The most effectively predicted is real estate (RMSE = 207) which, as can be seen in table 4.4, is also the sector with the smallest fraction of total flow value in point-to-point flows, at 14%. This relationship does not hold in general however, since the highest fraction, at 23%, is in the education sector (308), which is fifth-best estimated behind only health (265), and utilities and wholesale trade (both 282).

The most poorly estimated sector is hospitality (2570) which has only 15% of its dollar value in point-to-point flows. This, along with business services (2127) and other services (2084) have the highest RMSE by a very wide margin. This ordering from best- to worst-estimated is broadly consistent across all specifications. Recall that we are applying a model to services flows which was fitted on commodities flows. The persistent difference in how well services sectors are estimated by the fitted model perhaps

[6] The number of point-to-point records available for comparison is shown in brackets after the sector name. See table 4.4 for full details of the numbers of records in each sector.

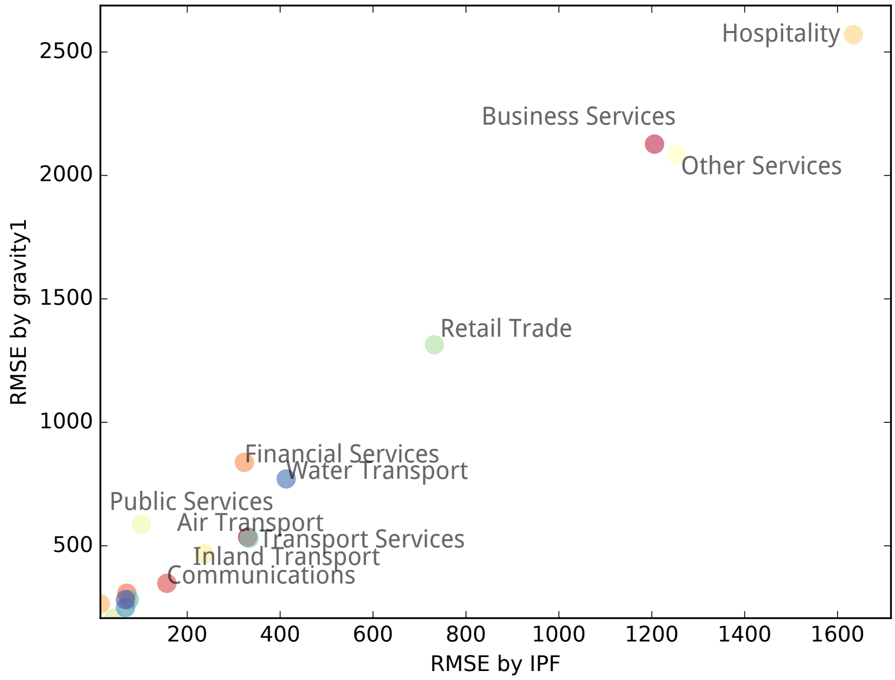


Figure 4.3: Root-mean-square estimation errors for each services sector from estimation via IPF and the *gravity1* specification. Sectors are labelled where practicable.

suggests that some sectors (namely real estate, health, utilities and wholesale trade) behave more like commodities sectors than do others (namely hospitality, business services, other services and retail trade.)

Specification (2) adds exporter final demand to (1) which improves the estimation for very few sectors and makes several very much worse (the exceptions being air transport, business services, hospitality and other services). The situation is somewhat improved by the addition of per unit value added in (3), but not completely salvaged until specification (4), which keeps value added but removes final demand. This latter is the best specification overall and, for some sectors, by a wide margin. The addition of country dummies in (5) does nothing to improve specification (3) to which it is identical bar the dummies, and therefore, for considerations of space, we do not test specification (4) in the presence of country dummies.

Recall that the parameters on these variables are taken from a regression run on commodity flows. The fact that the additional variables in specifications (2), (3) and (5) make the predictions worse than what we might term the baseline variable set indicates either that services flows have a fundamentally different relationship to the regressors, or that a gravity model is, in general, not doing a good job of estimating flow magnitudes.

In this context, it is interesting to see that the gravity model performs worse in every specification than IPF. Specification (4) has an RMSE of 11,851, making it the best of the gravity model specifications, but it is still a long way short of simple IPF 7,051. Certainly the unconstrained nature of the gravity model makes it naturally a poorer predictor of flow *magnitudes* than IPF, although it may do a better job of *ranking* the flows by magnitude. In future work, it would be interesting to test this hypothesis. But we can here present a novel method of combining the rankings of the gravity model with magnitudes of IPF.

4.4.4 Combining the gravity model and IPF

The results of the gravity model do not use the row and column totals which, in this context, are known. This is a considerable waste of the available information. We might therefore usefully use the outcome of the (unconstrained) gravity model as the input to an iterative proportional fitting routine. Recall from section 4.2.1 that flows which are larger in the initial matrix tend, in general to be larger in the final fitted result. Thus, the output of the gravity model provides relative differences in point-to-point flow values and the subsequent IPF ensures that the row and column margins balance while retaining "some" of the information from the first stage.

If this turns out to be more efficient than IPF alone, it would confirm our suspicions in section 4.4.3 that the gravity model is better than IPF at estimating the relative sizes of the flows, but poor at estimating the magnitude. This is simply because IPF takes care of the magnitude but will broadly retain the relative flow sizes from the gravity model.

We can think of this two-stage procedure as being an "operationalised" version of the doubly-constrained spatial interaction model of Wilson (1967) where, rather than explicitly calculating the balancing factors, A_i and B_j, we are simply operating on the flows predicted from the main section of that model's equation, such that the constraints are met.

The results of this two-stage operation are shown in table 4.3 with the columns labelled 'gipf' (for 'gravity IPF'). Here we see that the 5 specifications of the gravity model are almost identical, suggesting that the IPF has "smoothed out" the variation in the flow magnitudes of the pure gravity approach. But the important result is that these 'gipf' specifications are all far more effective than either gravity alone or IPF alone, in many sectors by a very wide margin. Business services sticks out as a particularly strong example of this.

This combining of a gravity model for the relative magnitudes, and IPF for the row/-column constraints is a novel method of improving model fit. An interesting direction for further work would be to apply this method to other scenarios, checking the results in-sample as we have done here.

4.5 Results

We will now turn our attention to an analysis of particular flows, comparing estimated values with values from data.

4.5.1 Selecting a representative sector

Since all the methods outlined above take each sector as a separate network, it makes sense to pick a sector for the analysis which has either the largest proportion of point-to-point flows to check against or, alternatively, the largest proportional dollar value of point-to-point flows.

Table 4.4 shows the situation for each of the services sectors. The first group of three columns shows the number of records against each sector. The first column shows the number of records which are point-to-point ("P2P"), which means that they are recorded as being both from and to a *country*, in the sense of an entity with a three-letter ISO code[7]. The inverse of this is "Regional" flows, shown in the second column. These are flows recorded as being either from, to, or in some cases from *and* to, non-country regions. These include geographical areas (such as "Eastern Europe n.e.s"[8]), trade areas (such as "Southern African Customs Union") and catch-all unknown categories (such as the very common "World" and the ultimately generic "Areas n.e.s."). The third column shows the fraction of all records for that sector which are point-to-point ("Frac. P2P").

Interestingly, the range of P2P fractions is very small, from 79% to 88%. It is also interesting that the fractions of P2P flows are all so high by this measure. This is presumably because the regional flows are large and monolithic, with flows recorded against large otherwise-unspecified regions. Because of this, we turn to the total dollar value of those flows. As before, we want to choose the sector with the largest proportion of dollar value contained within P2P flows.

Columns 4, 5 and 6 show the same measures as the first group but for the dollar value, measured in $US billions. Here again, the range is fairly small, just nine percentage points, but, this time, the fraction contained in P2P flows is much smaller, between 14% and 23%. Since the sector with the highest proportion of P2P flows by this measure (education) is comparatively small in magnitude, we might compromise between a large proportion of P2P and a large volume of trade overall. "Other Services" would fit this bill, but is a disappointingly vague choice for our representative sector. An ad-

[7] Examples of three-letter ISO codes are GBR for Great Britain and DEU for Germany.

[8] Where n.e.s stands for "not elsewhere specified"

Table 4.4: The services sectors categorised according to two measures. The first is the number of records in each category and the second is the total dollar value of all the flows in each category. In both cases, the two categories are point-to-point and regional, where regional is defined as any flow which originates or ends at a region, such as "Europe" or "World", rather than a country.

	Number of records			$ Value (billions)		
Sector	P2P	Regional	Frac. P2P	P2P	Regional	Frac. P2P
Education	1,382	184	0.88	31	106	0.23
Other Services	3,092	404	0.88	744	2,965	0.20
Construction	1,674	302	0.85	98	398	0.20
Financial Services	2,623	401	0.87	300	1,258	0.19
Business Services	3,421	405	0.89	997	4,498	0.18
Water Transport	2,976	424	0.88	449	2,159	0.17
Vehicle Trade	2,790	422	0.87	88	432	0.17
Utilities	2,791	423	0.87	94	458	0.17
Retail Trade	2,794	280	0.91	546	2,737	0.17
Air Transport	3,085	429	0.88	348	1,697	0.17
Inland Transport	3,104	428	0.88	248	1,217	0.17
Communications	2,586	372	0.87	107	507	0.17
Wholesale Trade	2,790	422	0.87	88	432	0.17
Transport Services	2,290	224	0.91	241	1,229	0.16
Health	896	163	0.85	5	27	0.16
Hospitality	2,914	430	0.87	1,211	6,619	0.15
Public Services	1,700	391	0.81	64	380	0.14
Real Estate	2,148	222	0.91	27	165	0.14

Table 4.5: The top ten most well-estimated and most poorly-estimated financial services flows using gravity model specification (4) followed by IPF. Error shown is simply the difference between the estimated flow and the flow in data. Only flows greater than $10M are included.

(a) The 10 best estimated

		Trade Val. ($US M)	Error ($US M)
ESP	CZE	10.12	0.13
BGR	AUT	10.32	0.15
ESP	SWE	10.65	0.23
DNK	BEL	16.55	−0.29
EST	DEU	13.21	0.37
AUT	FIN	10.71	0.43
AUT	TUR	13.65	−0.67
DNK	AUT	10.00	−0.80
ITA	IDN	12.73	0.99
BEL	CZE	23.40	−1.09

(b) The 10 worst estimated

		Trade Val. ($US M)	Error ($US M)
USA	GBR	15,574	−5,778
GBR	USA	15,668	−3,659
GBR	LUX	1,590	3,027
USA	JPN	4,261	2,752
ITA	IRL	3,265	−2,435
DEU	USA	6,907	−2,252
GBR	JPN	2,677	−2,218
FRA	GBR	1,508	2,189
FRA	LUX	3,538	−2,173
LUX	ITA	2,931	−1,956

ditional requirement is that it should be easy to interpret what trade in the services of that sector actually are. We therefore select financial services as a good compromise between these three competing requirements. It is a comparatively large sector, both in number terms and dollar value terms and is also comparatively easy to interpret as a traded sector (unlike, say, communications, whose interpretation as a traded sector would require significant additional interpretation as to what is and is not included.) Therefore the rest of this analysis will be performed on the financial services sector.

4.5.2 Estimated in-sample flows

The flow data for the financial services sector contains 2,623 P2P flows (those where the origin *and* destination are known). We can use these to test the effectiveness of each of the estimation methods outlined above by checking their values against the values estimated by each method where, of course, the methods have not had access to these values, only the row and column totals.

Table 4.5a shows the ten most well-estimated financial services flows using the "gipf4" specification. The error is simply the difference in dollar terms between the flow in the data, and the estimated flow. Flows smaller than $US 10 million are excluded. Perhaps predictably, since the errors are reported in absolute, not relative, terms, the best estimated flows tend to be smaller ones. This can be seen by comparing with the worst estimated flows in table 4.5b.

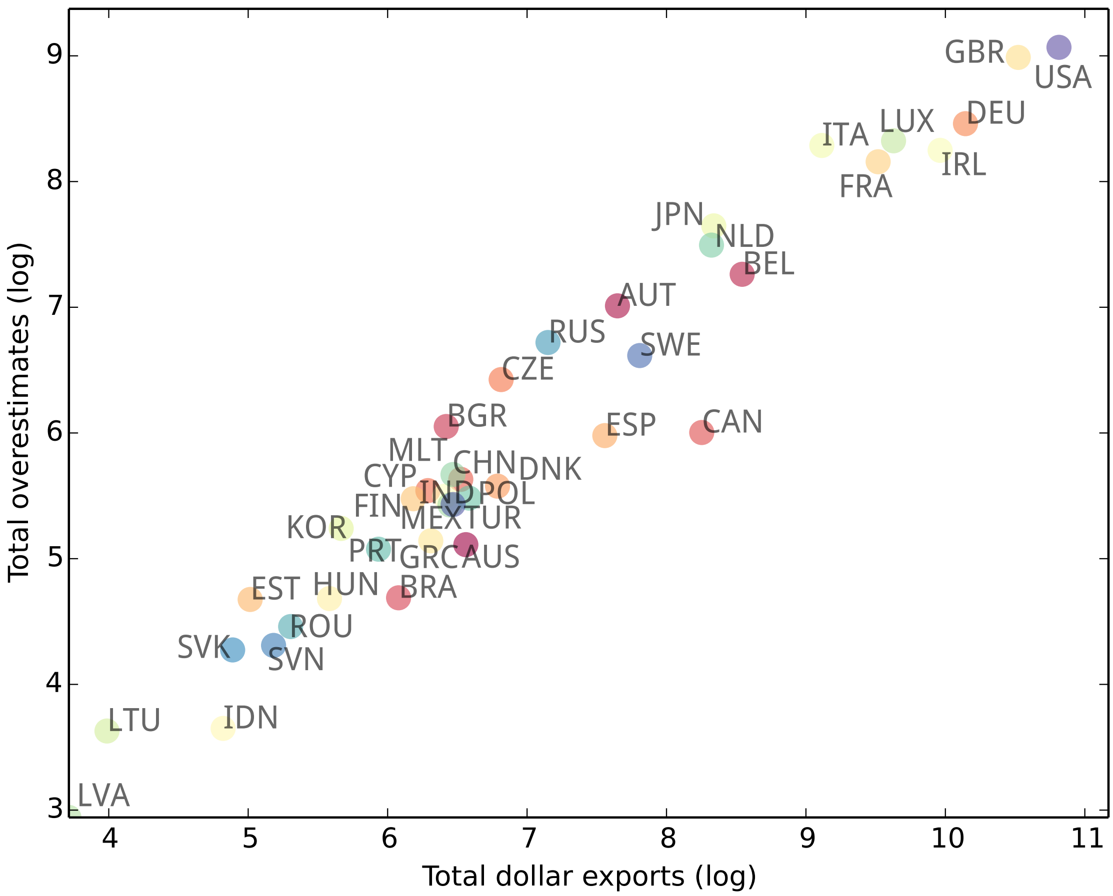


Figure 4.4: The total *gipf4* overestimate (the sum of all errors greater than zero) for each exporter, plotted against that country's total exports.

4.5.3 Estimated export totals

To have some faith in the estimation procedure, it would be reassuring to see that no country is consistently over- or underestimated. If this was the case, we might have reason to suspect that some important variables are being omitted in the specification used at the gravity stage of the estimation in section 4.4.4. In section 4.5.2, we saw that the estimation error seemed to be broadly related to the magnitude of the flow itself.

We might similarly expect that an exporting country's total estimation error (that is, the sum of the estimation errors across all flows emanating from that country) is correlated with the total exports of that country. If there are marked deviations from any such relationship between total error and total export, this might be evidence of a systematic bias in the estimation process.

Figure 4.4 shows a scatter plot of logged total exports against logged total *over*-estimation, the latter defined as the sum of all estimation errors greater than zero. The relationship is clearly linear, and is remarkably unscattered. This gives us additional evidence that the *gipf4* estimation process, while of course not being precise, is at least treating each country equally in terms of overestimation. The only country which might be considered an outlier is Canada, which is slightly less overestimated than its total export size would suggest, given the trend with all other countries. Visually speaking, though, this is not a hugely noticeable outlier.

If countries are generally overestimated to a similar extent, once we take into account their export size, we might be interested in whether the extent to which a country's underestimation is closely related to its overestimation. Again, if some countries are more likely to be under- than overestimated, this might suggest a bias in our estimation procedure. Table 4.6 shows a summary of the relationship between overestimation and underestimation. The countries are presented in order of absolute total error (ATE), defined as the absolute value of total overestimate minus total underestimate. The total under- and overestimates are also shown, to the nearest $US million.

It is striking how closely the under- and overestimates are balanced. This result is not seen in using IPF alone, where the total errors are far from zero in most cases. It seems that a combination of gravity model and IPF results in an extremely even balance between under- and over-estimation.

The most consistently underestimated country in terms of financial services flows is the USA, with a total underestimation of just $1,000.

We have shown that the *gipf4* specification is neither biased towards a particular country, nor towards under- or overestimation. We will therefore use *gipf4* in the estimation of point-to-point services flows throughout the remainder of the analyses which follow in this thesis.

Table 4.6: The five most and five least balanced exporting countries in terms of overestimation vs. underestimation. All figures are in $millions unless otherwise specified, and relate to the financial services sector. The errors are from the *gipf4* specification. ATE stands for absolute total error.

Country	Total export	Overestimate	Underestimate	ATE ($)
USA	49,745	8,671	−8,671	1001
GBR	37,146	8,002	−8,002	396
IRL	21,213	3,820	−3,820	110
DEU	25,441	4,724	−4,724	95
CAN	3,838	405	−405	73
⋮	⋮	⋮	⋮	⋮
SVK	133	72	−72	0.52
IND	608	244	−244	0.48
IDN	124	38	−38	0.31
LVA	41	19	−19	0.26
LTU	54	38	−38	0.24

4.6 Conclusion

In this chapter, we have shown how to deal with the fact that, in many cases, services flows are given only as import or export totals, not as point-to-point flows. We introduced the iterative proportional fitting (IPF) method as a "no information" simplest case. We showed that IPF is zero-retaining and broadly rank preserving.

We then went on to show how commodities flows, which are far better recorded and far more often point-to-point, can be used to better estimate services flows. To do this we introduced a generalised gravity model specification, which we then estimated in several specifications. In doing this, we uncovered the surprising fact that value added per unit has a negative exponent in the multiplicative gravity model which seeks to explain trade. We investigated this puzzle further by splitting value added into its constituent components and tentatively proposed that the use of primary products such as metals and minerals may be an active discouragement to trade, but that use of secondary products such as plastics, machinery and chemicals may encourage it.

Turning back to services flows, we performed an in-sample comparison of the estimates obtained by each specification of the gravity model and chose an optimum specification. We additionally found that IPF performed better than any of the gravity specifications and hypothesised that a combination of these approaches might be best.

We showed that this hypothesis was indeed true, and introduced the *gipf* (for gravity IPF) approach as an operationalised version of the doubly-constrained spatial interaction model. This was shown to be the best in terms of in-sample comparison and was also shown to be unbiased in its estimation of each exporter. The error was very closely correlated with the total export of a given exporter. Finally, the *gipf* specification was shown to produce almost exactly as much underestimation as overestimation, a result which further encourages its use in estimating point-to-point services flows from trade totals which we will proceed to do in every analysis which follows in this thesis.

Chapter 5

Analysis of the estimated extension to the model

5.1 Introduction

The WIOD data from which the demonstration model draws its country descriptions covers 40 countries, which are mainly European and mainly OECD members. The addition of other countries to the model requires either an additional data source or the numerical estimation of input-output tables (IOTs) for the other 150-odd countries of the world.

The work of Thomas Oléron Evans takes the latter approach, and uses flows into and out of the country to be estimated, which are often very well covered by the UN trade data, to estimate the intermediate flows which make up an input-output table.

The method uses machine learning, with the 40 countries of WIOD as a training set. Using this training set, the algorithm is 'taught' the relationship between the observed flows into and out of the country and that country's IOT, and then uses this information to estimate an IOT given an (assumed to be complete) set of flows into and out of each non-WIOD country. The method is based on an original idea by the author of this thesis, but the development of the idea and all of the results of the estimation itself are due entirely to work by Dr. Oléron Evans.

In this chapter, we will simply analyse the results of this numerical estimation in an attempt to demonstrate that the estimates produced are, if not a true reflection of an estimated countries economic structure, at least plausibly "IOT-esque" in their structure. This seemingly weak condition is more convincing in light of the fact that the IOTs estimated will in fact, and by definition, be able to reproduce the observed in- and out-flows of a country, as well as that country's GDP. It is hopefully clear that any IOT which “looks like” the IOTs in the training set, and is able to reproduce flows in and out of a country and that country’s overall economic size, is as good a replacement for an empirically observed IOT as we are likely to find.

5.2 In-sample analysis

We begin with an analysis of the IOTs produced by the machine learning algorithm for countries within the training set itself; a so-called "in-sample" assessment.

Each estimate produced represents either an intermediate flow from one sector to another within a particular country[1] or a final demand flow within that country. The estimation process "estimated" the 40 countries of the WIOD as well as the other 153 non-WIOD countries. This allows us to make a simple analysis of the estimates of the 40 countries with their "true" values as reported in WIOD.

5.2.1 Summary statistics

A summary of the results of this analysis are shown in figure 5.1. The WIOD value for each sector-to-sector flow or sector-to-final-demand flow (hereafter, we will simply call these "flows") is plotted against the equivalent estimated flow, with both axes being logged.

If the estimates were a perfect reconstruction of the data, we would expect every point to lie along the diagonal line, $y = x$. Any deviation from this represents a deviation of the estimated flow from the flow in data. Since this is a log-log plot, deviations at the right-hand side (larger range) are orders of magnitude more severe than deviations at the left-hand side (smaller range). It is therefore encouraging to see that as we reach the higher magnitudes, the points do seem to lie fairly close to the $y = x$ line and, in general, the plot shows a broad balance between under- and over-estimation with a slight bias towards over-estimation at the smaller scales of magnitude.

We can look in more detail at the estimates by breaking down the estimation error (estimated value minus the original data value) by country and by "from-sector" (i.e. the sector from which the flow originates.) Figure 5.2 shows summary statistics, broken down by both country (figure 5.2a) and from-sector (figure 5.2b). Estimation error is measured in the same units as the flows themselves, $US millions.

In each case, the elements of the graphs are as follows: the box represents the interquartile range (IQR) of the estimation error, the 25th percentile to the 75th percentile. The "whiskers" reach from the 10th to the 90th percentiles. The mean is represented by a small filled-in box and the median is an unbroken horizontal line.

[1] Within a particular year, but since all estimates are from 2005 this will be implicit in everything which follows. An interesting direction for future research could be to repeat the analysis of this chapter across the 16 years for which WIOD has data.

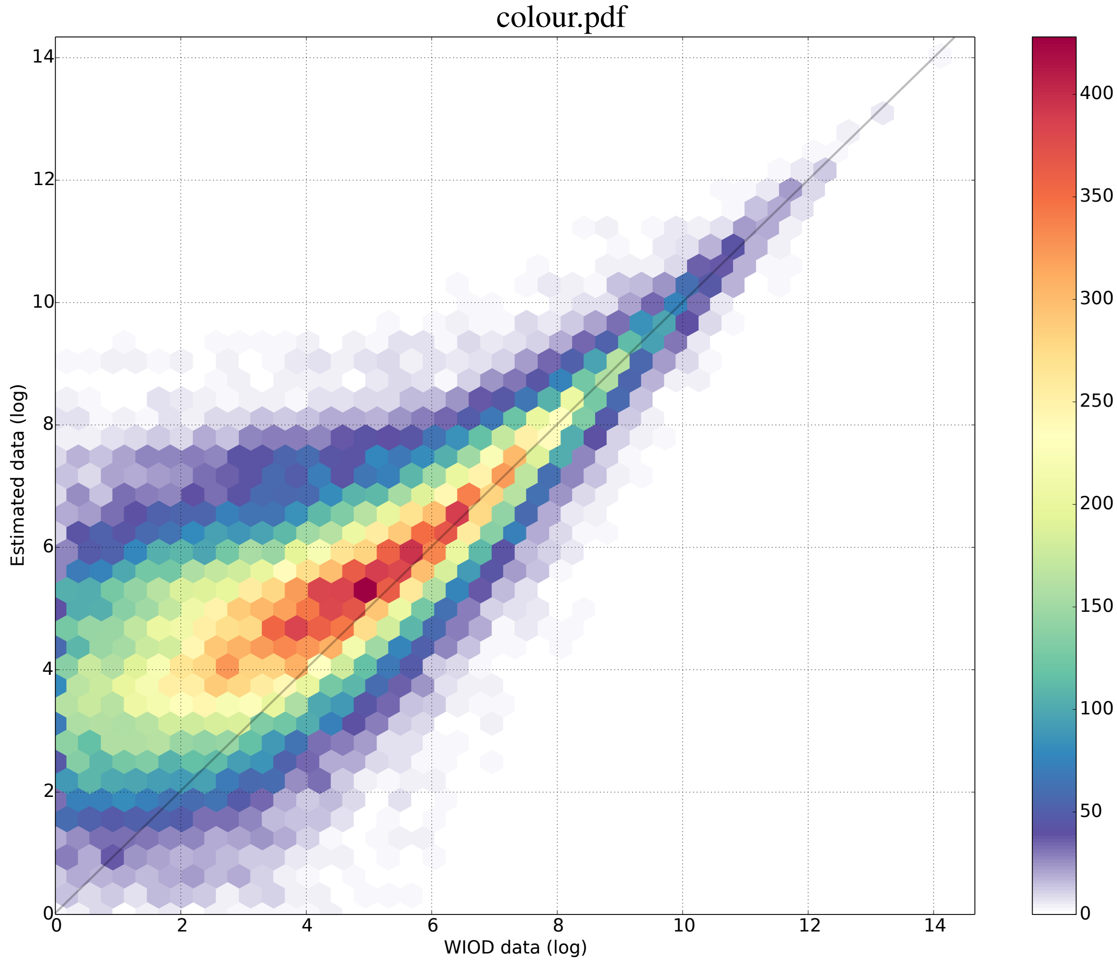


Figure 5.1: A density plot showing the relationship between the estimates of the 40 WIOD countries and the WIOD data itself. Different shades indicate a measure of the number of data points in each hexagon. All data points are flows, measured in $US millions. A little over 100,000 flows are plotted.

Figure 5.2a shows summary statistics for the ten biggest countries by total flow magnitude in the data. There is a general tendency towards over-estimation (the IQR strays into negative territory only for the USA), but this tendency seems to be exacerbated by very over-estimated "outliers". To see this, observe that the mean across all errors lies within the IQR of only three countries: Korea, Spain and the USA. In all countries, the median error lies relatively close to zero. The broadest 10th-90th percentile range of errors are those of the USA ($4.7 billion), Germany ($2.5 billion) and Japan ($1.9 billion), leading us to suspect that the estimation errors are simply a function of total GDP: that the spread of estimation errors is largest when the flows are largest overall. But the smallest are Korea ($680 million), Spain ($720 million) and China ($960 million), this last example showing that estimation error is not simply a function of total GDP.

On the from-sector breakdown in figure 5.2b, which as before shows the 10 biggest sectors by total flow magnitude, it is clear that much of the estimation error is being driven by the business services sector. That the most error-prone sector is a service, rather than a commodity, is to be expected since the data on services trade is far more poorly recorded than that on commodities trade (see section 4.1 for more details.) As with the by-country analysis, we see narrow IQRs with what seem to be a small number of very under- or overestimated flows, with means appearing far outside the IQR in four of the 10 biggest sectors, and in the case of construction, health and public services, outside the 10-90 decile range.

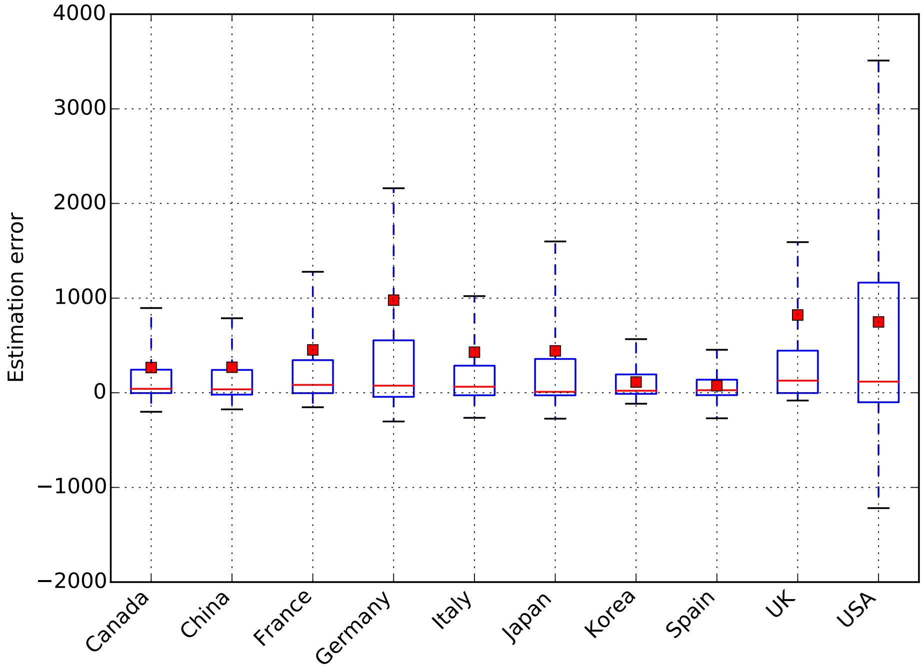
We will now look at the particulars of the estimation results in more detail.

5.2.2 Comparing estimated IO tables with heatmaps

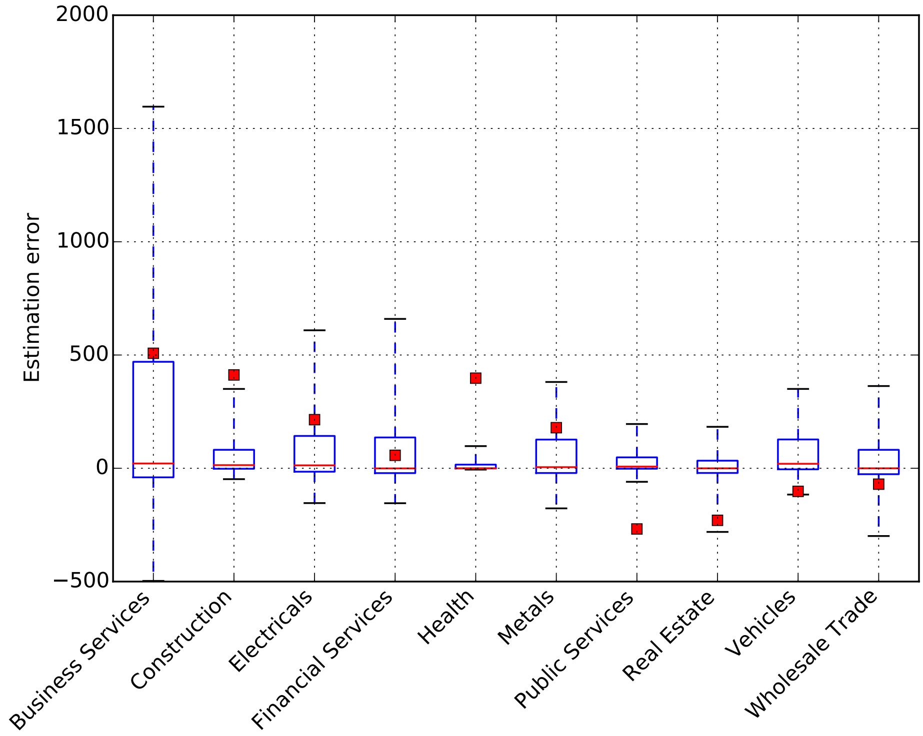
In order to look in more detail at some of the best- and worst-estimated countries from section 5.2.1, we can show the input-output table for a particular country as a heatmap. This involves a grid of sector-sector flows where the colour of the grid element is related to the size of the flow. This allows us a visual overview of the "shape" of a particular input-output table, and allows us to visually compare, albeit in a necessarily broad brush fashion, the WIOD input-output tables against the equivalent estimated tables.

We will restrict ourselves to comparing two well-estimated countries from section 5.2.1, Spain and China, and two more poorly-estimated countries, Germany and the USA. For visual clarity, we will also roll the 35 WIOD sectors into a more manageable seven "super-sectors". This is done by simple inspection of the sector names.[2] The

[2] The decision of which sectors to put into which super-sector is fairly clear cut, but the number of super-sectors is inevitably an arbitrary decision. This choice will of course affect the conclusions here. The somewhat unscientific nature of the categorisation should not be overly problematic though



(a) By country



(b) By "from-sector"

Figure 5.2: Boxplots showing the estimation error of flows within the 10 largest countries and the ten largest sectors by total flow magnitude. The dotted vertical lines show the 10th to the 90th percentiles. The box shows the interquartile range. The mean is represented as a small filled square and the median by a horizontal line. Estimation errors are measured in millions of $US.

Table 5.1: The seven "super-sectors" into which each of the 35 WIOD sectors has been aggregated in order to facilitate visual analysis of the estimates.

transport Inland Transport, Water Transport, Air Transport, Transport Services
primary Agriculture, Food, Leather, Paper, Textiles
secondary Chemicals, Electricals, Plastics, Manufacturing, Machinery, Vehicles
raw Wood, Mining, Metals, Fuel, Minerals
trade Vehicle Trade, Wholesale Trade, Retail Trade, Real Estate
services Business Services, Financial Services, Hospitality, Private Households
public Education, Health, Other Services, Communications, Construction, Public Services, Utilities

super-sectors are comprised from WIOD sectors as shown in table 5.1. For the sake of brevity, we will refer throughout to these super-sectors by name, *viz* "primary", and not by the more cumbersome but more grammatical "the primary sector".

Figure 5.3 shows heatmaps for the four countries mentioned, the USA, Germany, Spain and China and has the WIOD data input-output tables side-by-side with the equivalent set of estimates. Flows are from the row sector to the column sector. Note that flows to final demand are not included in these images. It should also be noted that each heatmap has its own colour scale, which is not shown. Thus, this exercise tells us only about the *shape* of each input-ouput table, not the absolute magnitudes of the flows inside them.

An encouraging initial observation is that the estimates are visually broadly similar to their data counterparts. We can therefore conclude that the estimation process produces what we might term "broadly realistic" input-output tables, by which we mean that the tables "look" like real input-output tables, even if idiosyncratic details per country are inevitably excluded. This was a key criterion mentioned in the introduction to this chapter.

Also encouraging is the fact that it is clear visually which estimate map goes with which WIOD data map: although the estimates and the data maps are not identical there is, at least among these four countries, no danger of mistakenly attributing an estimate set to the wrong data set. This is an important outcome, because the machine learning algorithm is dependent to a great extent on there being a balance between similarity among countries (in that there exists a "typical" input-output table to which we hope are estimates are reasonably close) and variation between countries to ensure that the outputs are sensitive enough to the inputs that all the estimated countries do not come out identical to one another.

as we will be using it primarily as a means for visual inspection. The categorisation is, though, reused later in this thesis in section 10.4.3 where the nature of the categorisation should be borne in mind.

Visually, the most well-estimated country is China, a finding which agrees with the summary statistic finding in section 5.2.1. In fact the estimate map is almost indistinguishable from the data map. Flows to services and trade from services and secondary are slightly under-estimated, but the estimation successfully finds that primary, secondary and raw are the main sectors in China along with flows to (but not from) public.

The other three countries have at least one visually obvious overestimate. In Spain, the overestimates lie along the diagonal, particularly secondary-secondary, and in Germany and the USA it is public-public and, to a lesser extent, services-public. The flows from services are generally slightly overestimated in both Spain and Germany, and flows from transport are overestimated generally in Germany and the USA.

Generally speaking, though, the shapes of the maps match surprisingly well. In Spain, the right half of the map, flows to services, trade and transport, is reproduced almost exactly, which agrees with our assessment in section 5.2.1 that Spain was among the most well-estimated countries.

With the exception of the overestimations mentioned above, the USA and Germany are also remarkably well estimated. This shows that, while the magnitudes of the flows in these countries may not always be well estimated, as we found in section 5.2.1, the overall cardinal nature, what we might call the “shape”, of the table may still be. This is a particularly important finding in this context, since the flows will largely be normalised via their conversion to technical coefficients in the usual input-output manner. It is therefore the shape, far more than the magnitudes, which is of importance.

On this basis, then, we conclude that the in-sample comparison of the estimates with WIOD figures is encouraging for the particular purpose at hand. A major risk with these tables is that they represent an over-fitting of the WIOD data: estimates match the data because the data have been used in the fitting. A useful validation exercise for future research would be to use what is called “cross-validation” whereby each data point (country) is removed from the fitting step, then introduced as a test case. This would give reassurance that the method is finding real information in the input-output tables, and not just reproducing what it has been given. But given our purposes here which is just to create broadly ‘realistic’ input-output tables which fit some key constraints, this extension is beyond the scope of this thesis.

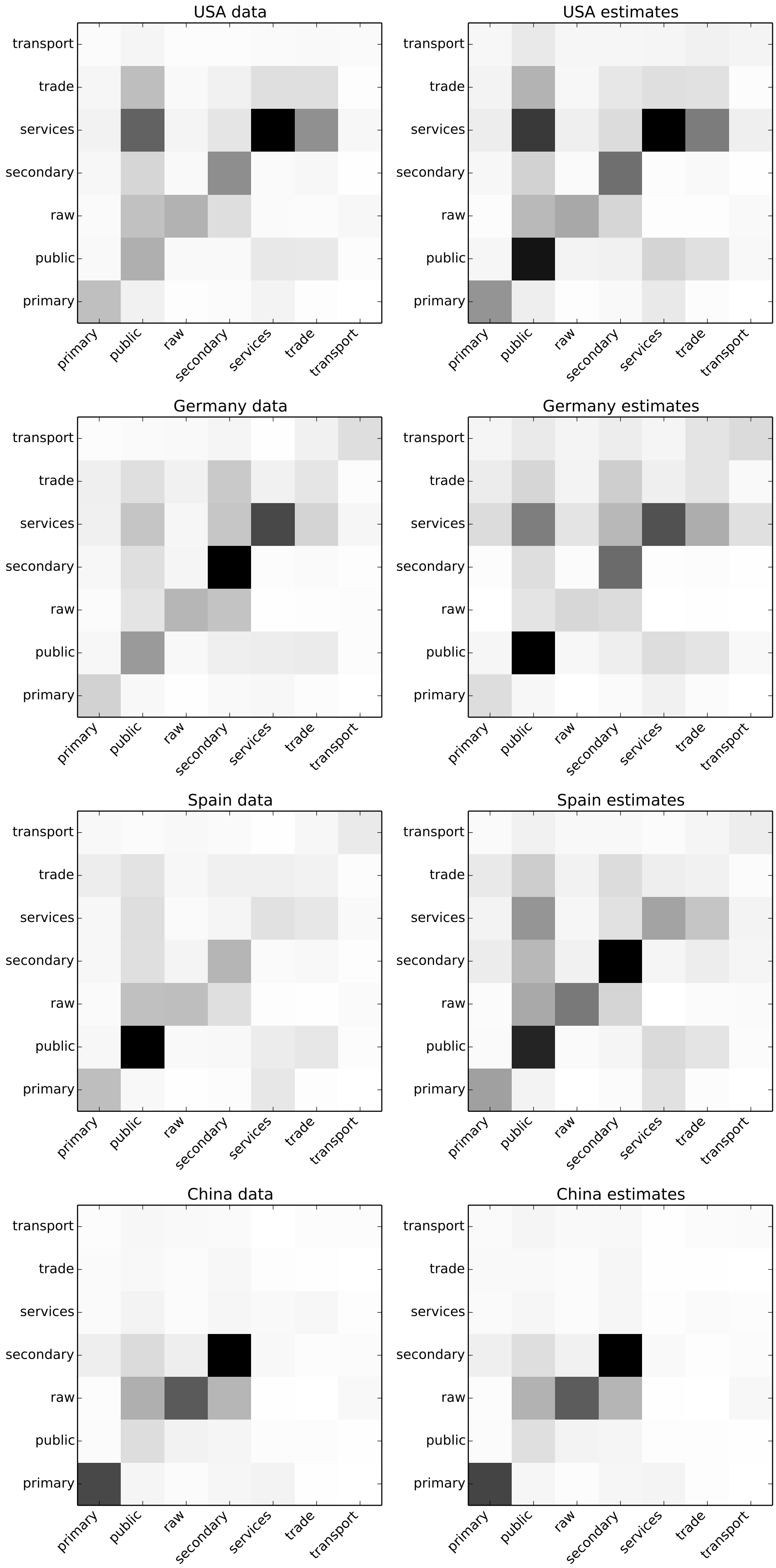


Figure 5.3: Visual comparison of WIOD data vs estimates for four countries.

5.3 Out-of-sample analysis

The discussion in section 5.2 has hopefully gone some way towards convincing us that the estimation algorithm estimates reasonably well those countries included in its training set, in cardinal terms (what we referred to as the "shape" of the resulting input-output table) if not always in ordinal terms.

But our purpose in running the algorithm in the first place is of course not to estimate the countries we already know, but to estimate those we do not. Beyond the fact that the estimation procedure produces input-output tables which, in some way, "look" like real input-output tables, rating the validity of the countries estimated out-of-sample is a more difficult task and will inevitably involve subjectivity in our assessment of the quality of the estimates.

We will start by developing a simple distance measure which will allow us compare countries to one another. In this way we will be able to see which WIOD country an estimated country most resembles and see if the results are in any way intuitively appealing.

5.3.1 Final demand closeness

We introduce a method of categorising estimated countries based on the estimated final demand for domestic goods and comparing each to the same set of final demands in WIOD countries. For simplicity of exposition, we will only examine countries with populations of more than 10 million. This cut-off leads to the inclusion of 52 countries in the comparison, a similar number to the 40 WIOD countries.

We begin by treating the final demand for domestic goods as a vector, defined as follows:

$$\mathbf{f}_i = \begin{bmatrix} f_{i1} \\ f_{i2} \\ \vdots \\ f_{iS} \end{bmatrix} \tag{5.1}$$

where f_{is} is the final demand for domestically-produced sector s and S is the total number of sectors. This then allows us to think of the direction in which each of these S-vectors "points", and of the concept of there existing an angle between any two such vectors.

Table 5.2: The 52 non-WIOD countries with a population of greater than 10 million, grouped by the WIOD country to which the final demand vector is closest in angle, defined by equation (5.2).

WIOD country	Estimated countries
China	Philippines, Thailand
Estonia	Argentina, Colombia, South Africa
India	Bangladesh, Cambodia, Sri Lanka, Morocco, Tunisia, Vietnam
Indonesia	Afghanistan, Angola, Burkina Faso, Ivory Coast, Cameroon, DRC, Ghana, Kazakhstan, Madagascar, Mali, Mozambique, Malaysia, Nepal, Sudan, Senegal, Tanzania, Uganda, Uzbekistan, Yemen
Lithuania	Chile
Mexico	Iraq, Malawi, Niger, Chad, Ukraine, Zambia, Zimbabwe
Russia	Algeria, Iran, Nigeria, Saudi Arabia, Venezuela
Turkey	Cuba, Ecuador, Egypt, Ethiopia, Guatemala, Kenya, Pakistan, Peru, Syria

By normalising the vectors, thereby setting their lengths to one, we can find the angle between any pair by using a dot product and some trigonometry:

$$\theta_{ij} = \arcsin\left(\bar{\mathbf{f}}_i \cdot \bar{\mathbf{f}}_j\right) \tag{5.2}$$

where a bar shows that a vector has been normalised.

This leads us naturally to a categorisation of the estimated countries based on which WIOD country has the closest final demand vector by this measure. Table 5.2 shows the results of this experiment for the 52 large-population countries, where each estimated country is placed with the WIOD country for which equation (5.2) is a minimum.

There are eight WIOD countries which are a non-WIOD country's closest neighbour: China, Estonia, India, Indonesia, Lithuania, Mexico, Russia and Turkey. All these countries are at the lower end of the GDP-per-capita spectrum, which might be considered encouraging for this exercise, since WIOD countries are generally richer, European and OECD countries.

Also encouraging is the fact that the two countries in the China group are the Philippines and Thailand, both south-east Asian countries. Bangladesh, Cambodia, Sri Lanka and Vietnam align with India in a similarly encouraging geographic fashion, although Pakistan does not. Malaysia aligns with Indonesia. Egypt, Pakistan and Syria align with Turkey. By this geographic reasoning, we might perhaps hope that the South American countries on the non-WIOD list would align to either Brazil or Mexico, but this is not the case.

But since these are *economic* descriptions, not geographic ones, it might be more useful to discuss similarities in economic terms. There is a set of oil-producing countries which align with Russia: Algeria, Iran, Nigeria, Saudi Arabia and Venezuela, all of which are members of OPEC. The only other OPEC members on this list are Angola and Iraq, who do not fit this pattern, aligning instead with Indonesia and Mexico respectively.

Argentina, South Africa and Chile are among the richest countries on the list and align with Estonia and Lithuania, the two richest of the WIOD category countries. We might think of these therefore as a high-GDP-per-capita group.

The largest group by far is that aligned with Indonesia, the poorest of the category countries after India. In this group are Afghanistan, Kazakhstan, Nepal and Uzbekistan along with all the sub-Saharan countries on the list apart from Malawi, Niger, Chad, Zambia and Zimbabwe which align with Mexico; and Ethiopia and Kenya which align with Turkey. We might therefore label the Indonesia group as the low-GDP-per-capita group.

In summary, despite the *post-hoc* nature of the rationalisation, we may find some mild source of encouragement in table 5.2, particularly where we have labelled a grouping above. But we also find several causes for puzzlement. The grouping around Mexico in particular is difficult to explain.

But two points should be borne in mind. Firstly, the way the experiment is devised, each estimated country is paired with exactly one WIOD country, and we have made no attempt to measure the closeness of this pairing. It may be that the Mexico group is simply a group of "leftovers" which are not particularly similar to any WIOD country. Secondly, that this experiment gives a precedence to the final demand vector (and only final demand for domestic goods) over all other aspects of the estimates, which the estimation process itself does not share. Thus, by focusing in this way on the final demand vector, we may be excluding further encouraging similarities with WIOD countries (or, of course, be excluding further puzzling alignments.) We deal with the second of these points in the next section with somewhat more convincing results.

5.3.2 Technical coefficient clustering

The analysis of section 5.3.1 grouped estimated countries with their "nearest" WIOD neighbour according to their final demand vectors. We might also ask how the estimated countries group to one another in terms of their input-output tables more broadly, and whether any geographic or macroeconomic justification might be found which will encourage us that the estimation process has done a reasonable job.

It is important to focus on the “shape” of the estimated input-output tables, as we did in section 5.2, rather the magnitude of any particular flow, and this brings us naturally to studying technical coefficients, which are not only unitless but are constrained to sum to less than unity over each column. In a similar fashion to the aggregation used in the demonstration model itself, we will aggregate the intermediate flows from domestic and imported goods, thus focusing more on the technical aspects of the production process and less on the significance of imports over domestic goods.

The first check of a sensible input-output table is that none of the columns of technical coefficients sum to greater than unity. In the case of this set of estimates, this is true in all cases except for four sectors in Bermuda. For this reason, Bermuda will be excluded from all the analysis which follows in this entire thesis.

To group the countries by the similarity of their technical coefficients, we will use the k-means clustering algorithm which is very simple, and implemented in many statistical packages. Although it is a well-understood procedure in computer science, it is still used rarely in economics.

One drawback the particular algorithm has over others is that it involves specifying a number of clusters *a priori* which is difficult to do with any rigour. Additionally, as with all clustering algorithms, k-means has an element of stochasticity in its output, with multiple runs of the algorithm producing different end results.

To deal with both these problems simultaneously, we will run k-means many times, and aggregate the results into a “co-occurrence” matrix. This is a square symmetric matrix with columns and rows indexed by the countries in the analysis. A cell ij records each time a particular run of the clustering algorithm places country i in the same cluster as country j. In this way the co-occurrence matrix represents a picture of the various runs of the clustering algorithm which we can visualise as a heat map with darker colours representing a more frequent pairing of the row and column countries in a cluster. The diagonal line represents the maximum co-occurrence, since a country always appears in the same cluster as itself. This is a further innovation in economic modelling of this kind.

A mixed picture emerges from figure 5.4, which shows the results of this experiment, with three well-defined clusters in the bottom-right corner of the matrix, and perhaps up to three more poorly-defined overlapping clusters in the top-left. The large cluster at the far bottom-right is what we might term the “low-income Africa cluster”, consisting of Burkina Faso, Mali, Malawi, Niger, DRC, Chad, Zimbabwe, Zambia and Madagascar. Also in this cluster, less neatly, are Nepal, Cambodia and Afghanistan. Senegal and

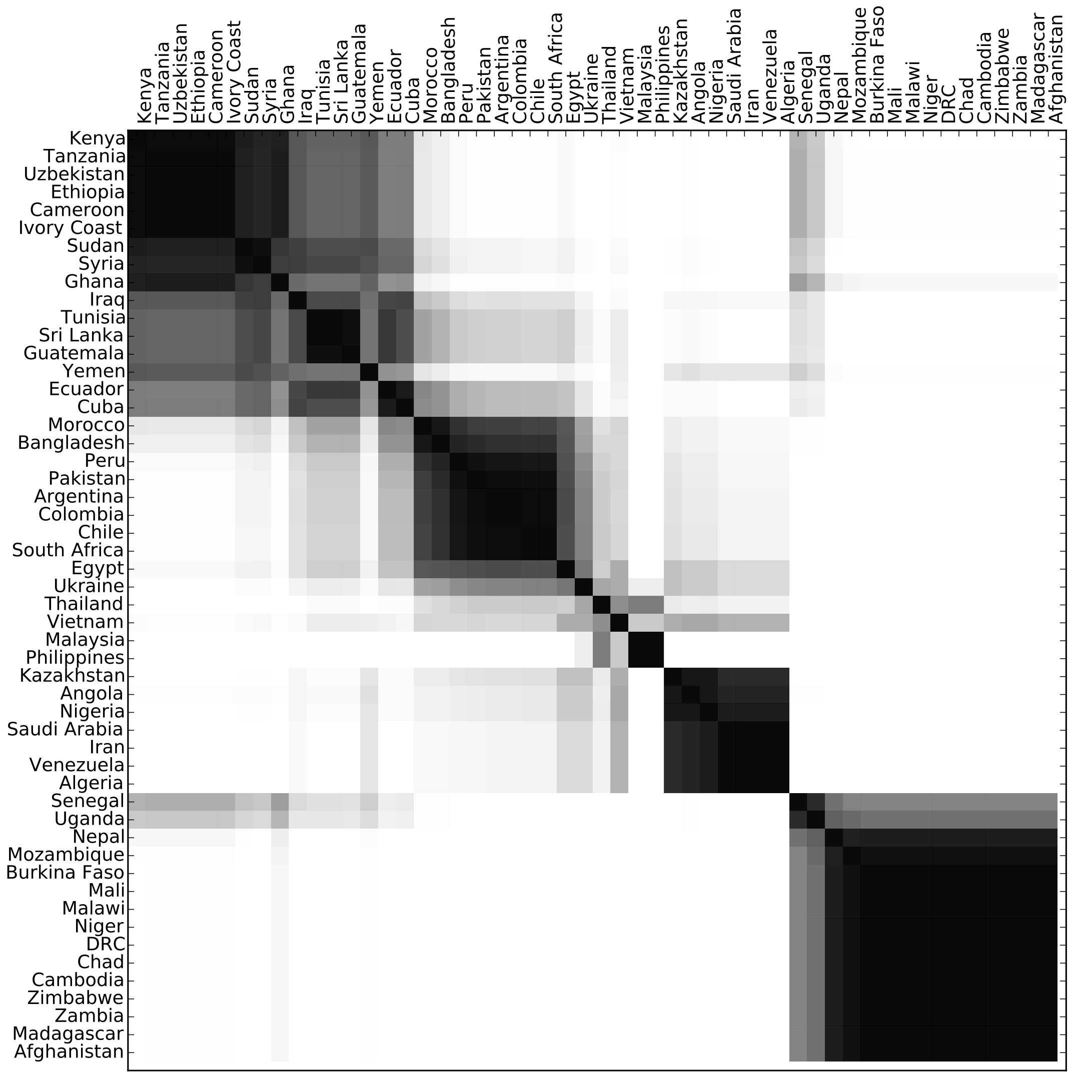


Figure 5.4: A co-occurrence matrix showing the result of 100 runs of the k-means clustering algorithm, set to split the technical coefficients of the countries shown into 5 clusters. Darker colours indicate a more frequent co-occurrence of the pair of countries in the same cluster.

Uganda are weakly in this cluster as well as weakly in the other major African cluster at the top left, but are most often placed in a cluster of their own.

The "OPEC cluster" we identified in section 5.3.1 is very well defined at second-from bottom-right, containing Kazakhstan, Angola, Nigeria, Saudia Arabia, Iran, Venezuela and Algeria. Vietnam is also weakly in this cluster, but is most often in a cluster of its own.

A "South-East Asia cluster" containing Malaysia, the Philippines and Thailand and, more weakly, Vietnam is the third of the well-defined clusters of the bottom-right. Visually, it is Vietnam's occasional membership of the OPEC cluster which prevents this cluster from being as sharply defined as the other two.

The picture in the top-left is more complex, and difficult to categorise. We have a strong "South American cluster" of Argentina, Columbia, Chile and Peru, but which also contains Bangladesh, Morocco and Egypt. Ecuador and Cuba also belong weakly to this cluster although more often appear as part of a large cluster in the top-left.

Another strong cluster of African countries at the top-left is made up of Kenya, Tanzania, Ethiopia, Cameroon, Ivory Coast and Ghana, tentatively a "high-income Africa cluster", although this cluster also strongly features Sudan as well as Uzbekistan and Syria. Within the larger weak cluster of the top-left is an entirely non-geographic cluster of Tunisia, Sri Lanka and Guatemala.

It is particularly intriguing that the countries of Africa are divided very sharply into two completely non-overlapping clusters with Senegal and Uganda the only "bridge countries". On one side, followed by numbers indicating their position in the World Bank richest country rankings are Burkina Faso (166), Chad(157), DRC (182), Mali (169), Malawi (183), Madagascar (173), Mozambique (179), Niger (180), Zambia (136) and Zimbabwe (163).[3] On the other side are Cameroon (149), Ethiopia (176), Ghana (134), Ivory Coast (142), Kenya (150), Sudan (140) and Tanzania (154). The two bridge countries are Senegal (156) and Uganda (167). It appears that the African countries are thus broadly split by GDP per capita into a richer cluster at the top-left and a poorer cluster at the bottom-right. By this categorisation, only Zambia and Ethiopia are in the "wrong" cluster.

This multiple clustering gives us a more shaded view of how estimated country input-output tables relate to, and differ from, one another. This section, as with the last,

[3] Richest countries measured in purchasing-power-parity adjusted GDP per capita for the years 2011 to 2014.

has shown some encouraging signs and some puzzling ones. Overall, it seems that the estimation process used has some merits, and needs some additional work.

5.4 Conclusion

In this chapter we have looked in some detail at the estimates produced by a machine-learning algorithm with observed import and export figures as inputs, and an input-output table (IOT) as an output, using the 40 countries of WIOD as a training set to predict IOTs for 152 non-WIOD countries. We focused on 52 of the largest non-WIOD countries by population.

First, looking within the 40 countries (in-sample), we found that the estimates were broadly of the correct magnitude, with a tendency towards over-estimation, particularly among the smaller magnitude flows. We then saw that much of the overestimation on a per-country basis was likely to be attributable to a small number of large mis-estimations, rather than a large number of smaller errors. We also found that the business services sector was a large contributor to overall mis-estimation.

We were encouraged by a visual comparison of in-sample estimates, both that the estimated IOTs were matchable to their empirical counterparts and that there was enough variation in the empirical IOTs to easily distinguish between them visually. We argued that these were both important aspects of the estimation process and we justified the use of the visual approach, and indeed were further encouraged, by the fact that the IOTs were to be converted to technical coefficients before use in the demonstration model, such that cardinal relationships were more important than ordinal ones.

Out of sample, we found a certain amount of geographic and economic evidence for the appropriateness of the method in creating plausible IOTs by clustering first by the final demand vector and second by the technical coefficients. We were able to tentatively distinguish between the IOTs of OPEC, poor African, rich African and south-east Asian countries. Other countries were less well distinguished and we were able to make little or no sense of the IOT-clustering of several of our sample of 52 large countries.

Nevertheless, we believe that this section gives us enough evidence that the IOTs produced by the machine learning estimation algorithm are sufficiently "like" genuine IOTs, and perhaps also broadly reflect the "category" of a particular country. They can therefore can be used in a modelling context whereby the GDP of each country and each country's trade relationship with the rest of the world are not drawn from the estimates but from UN data. This means that the ability of the estimates to affect outcomes is somewhat more limited that in an analysis where *only* the estimates were used.

It is left to the IOT merely to act in a plausible way in response to changes in external flows. We have presented some evidence here that the estimated IOTs are similar enough to real IOTs (albeit not necessarily the correct ones) that this somewhat weaker requirement may be fulfilled.

Part III

Analysis and Applications

Chapter 6

Analysis of the model itself[*]

6.1 Introduction

Now that we have assembled the data, mathematical infrastructure and metrics outlined in part II, we are in a position to demonstrate how some commonly-undertaken social science analyses can be given a different perspective by their inclusion in a global economic model.

We can adjust some or other of the final demand levels and the three sets of coefficients, technical coefficients, import ratios, and import propensities, in order that the resulting model matches various proposed scenarios. We can then measure the impact of these changes on the GDP and balance of trade (BoT) of the countries in the model.[1]

In this chapter we present some of the simplest possible analyses as a "first glimpse" of the kinds of output which can be produced by a model of the global economy such as we have now constructed. We can see these analyses and, to some extent, those that follow them, as a suggestion or guide of how to adapt whichever policy questions a future user of the model may have.

The analyses of this chapter are "within-model" in the sense that they employ only the coefficients of the model themselves as opposed to adding any data or assumptions external to the model. This is in distinction to the analyses of subsequent chapters which all add external material of one kind or another.

In this section, we present three within-model analyses. Each of them looks at one particular property of each sector in each modelled country and uses this to assess the

[*] This chapter uses the 40-country model throughout. See page 77 for a description of what this means.

[1] In fact, what we are calculating is total value added (TVA). But since this is one of the possible definitions of GDP, we will use the terms TVA and GDP interchangeably throughout everything which follows.

impact of each sector on the global economy. From here on, we will refer to a particular country/sector combination, for example German Vehicles, as a “CS”.

We start in section 6.2 with demand sensitivity, a measure of the response of the global economy to changes in final demand in a particular CS. Next, in section 6.3, we split demand sensitivity into domestic and foreign sensitivities and use these to say something about the economic self-sufficiency of each country. Finally, we look in section 6.4 at how the import ratios can be used to assess the effect of import substitution, the replacement of imported goods for those domestically produced, in each CS.

In all the analysis which follows, a model based on data from 2010 is used. All dollar measures are given in current price (i.e. 2010) US dollars

6.2 Demand sensitivity

The first of our analyses is perhaps the simplest possible analysis which can be performed using the demonstration model.[2] The experiment consists of the sensitivity of some global metric with respect to exogenous changes in final demand in each CS, what economists call a "demand shock".

A natural choice of metric might be "global GDP": a global/domestic contradiction-in-terms which nevertheless has an obvious meaning. But in a closed model where no production goes to waste and all demand is fulfilled, any value which is created must be consumed somewhere. Thus total value added (which, as outlined in section 3.4.5 is equivalent to GDP) is always equal to total final demand. Hence a reduction of, say, \$1 in the final demand of any CS will result in a reduction of precisely \$1 of global GDP.[3]

6.2.1 Gross global output

Instead of total value added, we instead use gross global output (GGO), defined as:

$$\text{GGO} = \sum_i \sum_s x_{is} \tag{6.1}$$

and measure the change in this metric with respect to a negative shock of \$1 million in f_{is}, the final demand of each CS. For clarity, we define the change in GGO in response to a reduction of final demand for sector s in country i as $\Delta_{is}\text{GGO}$. This is not only takes account of the *direct* effects of such a change (a reduction in the output of sector i and of those sectors supplying sector i's intermediate demand) but also of the full spectrum

[2] Indeed, intuition might suggest that it hardly needs to be carried out at all, since the model is entirely deterministic in nature and the coefficients are all known at the time we carry out the "experiment". But nevertheless, with a model the size of ours (each country has 1,225 technical coefficients, 35 import ratios and 1,400 import propensities) numerical approaches such as those presented here are likely to be more appropriate than in a smaller-scale system, where closed-form solutions to such questions can be stated. More mathematically satisfying methods of performing such experiments (albeit in far smaller systems) can be found in, e.g., Sierksma (1979). A far more complex version of this same analysis, including sensitivity to many aspects of an input–output model, is given in Chapter 5 of Chenery, Robinson, and Syrquin (1986). Additionally, this particular experiment turns out to reveal something interesting about the nature of the Chinese economy, as we will see in section 6.2.3.

[3] This fact serves as a stark reminder that we are dealing here with a model of the global economy, not with any recognisable "real" economic system. Economies of scale at the production and delivery end, frictions of every imaginable sort including delays in production and delays in stocking and ordering, and difficult-to-model factors such as consumer confidence mean that no real economy is a closed system with a one-to-one relationship between TVA and final demand. Rodrik (2015) talks eloquently and at length about the differences between models and the real world, and how models can be useful without being realistic. Nevertheless, this demonstration model is very simple indeed, as has been pointed out elsewhere, and future work should involve running these same analyses with a more complex CGE/GTAP model. However, the accuracy of the model used is not our main focus in this thesis.

Table 6.1: The 10 most significant countries, sectors and CSs in terms of their effect on gross global output (GGO) in response to a negative final demand shock of $1 million. All figures are shown in $millions. GGO figures in (a) and (b) are summed over either sector or country as appropriate.

(a) The 10 most significant countries

Country	Δ_iGGO
China	82.67
Czech Rep.	69.62
S. Korea	68.43
Russia	67.95
Bulgaria	67.36
Japan	66.76
Italy	66.52
Poland	66.39
Finland	66.03
Portugal	65.64

(b) The 10 most significant sectors

Sector	Δ_sGGO
Vehicles	98.09
Plastics	87.40
Food	87.05
Machinery	85.91
Metals	85.86
Leather	85.49
Wood	85.49
Electricals	85.41
Textiles	84.00
Construction	83.40

(c) The 10 most significant country/sectors

Country	Sector	Δ_{is}GGO
China	Vehicles	3.28
China	Textiles	3.22
China	Plastics	3.20
China	Leather	3.14
China	Construction	3.05
China	Wood	3.02
China	Paper	3.00
S. Korea	Vehicles	3.00
China	Machinery	2.97
China	Metals	2.96

of *indirect* effects: on those sectors which supplied the sectors which supplied sector i, on *their* suppliers and so on.

To recap, the experiment involves: 1) reducing the final demand of the first CS by $1 million, 2) measuring Δ_{11}GGO and 3) returning the model to its initial state before moving on to the next CS. We have found, in experiments whose results are not reported here, that due to the linearity of our model GGO responds precisely linearly to changes in final demand and therefore the second derivative will not be reported. Any CSs with an initial final demand of less than the magnitude of the shock will be ignored, to prevent final demand being set negative. Table 6.1 shows three different summaries of the results of the experiment.

6.2.2 Results

We focus first on the countries with the largest responses overall, measured in terms of the sum of the responses over each of the country's sectors:

$$\Delta_i \text{GGO} \equiv \sum_s \Delta_{is} \text{GGO} \tag{6.2}$$

If a demand shock in a sector induces a large reduction in production globally, it follows that sectors with a larger value of Δ_{is}GGO must require, once the entire production network is accounted for, a larger amount of input to produce each unit of their output. Δ_iGGO can thus be thought of as a country-level significance measure, with larger values implying a economy which, summed over the effects of each of its sectors, has a greater significance on production in the rest of the world. Table 6.1a shows the ten most significant countries by this measure.

GGO is most responsive to China's sectors, with \$83 million across all 35 sectors of the Chinese economy. This is perhaps a surprising result, as we might have expected the US to be the world's most significant economy. To investigate this further, we will in section 6.2.3 split the significance measure, Δ_i, into foreign and domestic effects, which will turn out to be revealing about the nature of the Chinese economy and help to explain China's significance according to this measure.

The nine countries which come below China are all fairly close to one another, ranging from \$70 million for the Czech Republic, to \$66 million for Portugal. Also among the ten countries to which GGO is most sensitive are South Korea, Russia, Japan and Italy. Indeed, although this is not shown in the table, the responses to all countries are fairly similar, with the least responsive being Luxembourg and Greece with \$56 million each. It seems that China is a real outlier, it being the only country with a significant gap between it and the countries near it in this ranking. We now take a closer look at why this might be the case.

6.2.3 Domestic vs. foreign effects

We begin by looking at the China result in more detail in table 6.1c, where the 10 CSs associated with the largest Δ_{is} are shown.

Nine of the ten most GGO-affecting CSs are in China. This result shows that something about China's economy is different to that of other countries across *all* sectors, although the most affecting is the Chinese vehicles sector, with \$3.3 million of GGO lost in response to a drop of \$1 million in demand.

We might be interested to know *where* this huge effect of Chinese vehicles, and of the Chinese economy more generally, is being felt. To investigate this, we split our Δ_i measure into two effects. The first is a domestic effect, looking at how changes in final demand in country i's sectors affect the gross output (GO) of country i. Formally:

$$\Delta_i^D \mathrm{GO} = \sum_s \Delta_{is} \mathrm{GO}_i \tag{6.3}$$

The second is a foreign effect, which looks at how those same country i changes affect GO in all other countries[4]:

$$\Delta_i^F \mathrm{GO} = \sum_{j \neq i} \sum_s \Delta_{is} \mathrm{GO}_j \tag{6.4}$$

Henceforth dropping the "GGO" and "GO" specifiers, we also notice the relationship to the term defined in equation (6.2):

$$\Delta_i = \Delta_i^F + \Delta_i^D \tag{6.5}$$

We have found Δ_i^D and Δ_i^F for every modelled country and plotted the two variables against one another on a scatter plot. The results are shown in figure 6.1 where a linear fit has been added to the plot.

There seems to be a very close systematic relationship between Δ_i^D and Δ_i^F, implying that most countries have a similar total significance Δ_i. This confirms the result we saw in table 6.1a. The region below the line is the region of less-than-average significance and that above the line is the region of greater-than-average significance.

Most noticeable about this plot is that China is an extreme outlier. Its economy is around one-third more significant ($\Delta_i = 82$) than that of Brazil ($\Delta_i = 61.2$) or India ($\Delta_i = 60.8$), but almost all of this extra significance relates to its impact on its own domestic sectors. Given China's Δ^F of around \$10M, we would "expect", according to the linear model which fits so well for all other countries, a Δ^D of around \$55M, but we in fact see the domestic response being closer to \$75M.

[4] We should note here that the 'rest of the world' (RoW) is not included in this analysis. Thus "foreign" here refers only to the countries explicitly modelled. According to Streicher and Stehrer (2012), the 40 WIOD countries cover 85% of the global economy, so the effect of excluding RoW should not be too large in magnitude, and will likely affect very little the comparative results within the countries in the model.

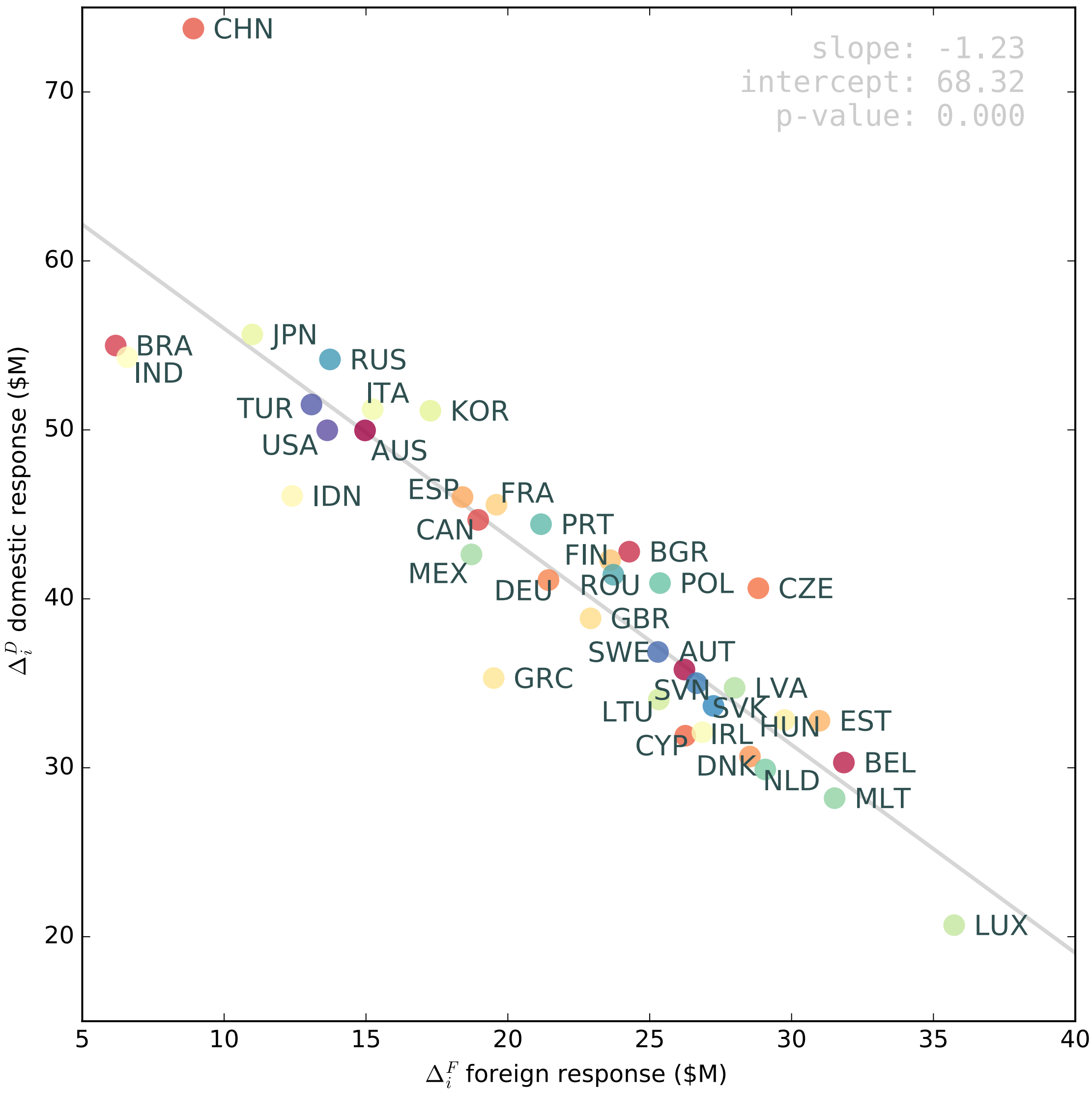


Figure 6.1: The total response of the global economy to a small negative final demand shock, summed across every sector in a country. 'Response' is defined as the total production lost across all sectors. Here, response has been divided into domestic, where the response is measured only in the country whose final demand has been reduced, and foreign, where the response is measured in all other countries.' The solid line shows the result of a linear fit.

6.2.4 Conclusions

We have shown that China's significance to the world economy originates largely in China itself: it is significant to the extent that it uses disproportionately large amounts of its own productive output in its production technology.

In an earlier study, Koopman, Wang, and Wei (2008) discuss in detail the extent to which domestic products contribute to the Chinese economy, in this case in terms of their presence in Chinese exports. We might also tentatively put this current result into the context of a discussion around the end of low labour costs in China (Li et al. 2012; Economist 2012). Future work might investigate the changing foreign/domestic significance of China over time.

6.3 Economic self-sufficiency

In countries further to the left of figure 6.1, such as Brazil, India, Japan, Russia, Turkey and the USA, a domestic demand shock prompts a comparatively smaller response in the rest of the world that it does within the countries themselves.

We could therefore describe these countries as being more self-sufficient than those on the right, such as Luxembourg, Belgium, the Netherlands and Denmark, in the sense that a greater percentage of the goods and services consumed in self-sufficient countries contain, as either primary or intermediate inputs, goods and services originating within those countries.

6.3.1 A measure of self-sufficiency

We can make this concept of self-sufficiency precise by measuring the ratio of domestic response to overall (foreign *and* domestic) response following a demand shock:

$$\phi_i = \frac{\Delta_i^D}{\Delta_i^D + \Delta_i^F} = \frac{\Delta_i^D}{\Delta_i} \tag{6.6}$$

such that a country with $\phi_i = 1$ is completely self-sufficient, and causes no impact to the wider world due to a demand shock, and one with $\phi_i = 0$ is entirely dependent on imports.

6.3.2 Results

By this measure, the most self-sufficient countries in the model are Brazil (0.90), China (0.89), India (0.89) and Japan (0.84), and the least self-sufficient are Luxembourg (0.37), Malta (0.47), Belgium (0.49) and the Netherlands (0.51). The United States is the 8th most self sufficient.

These results might appeal intuitively, but it might also seem that we are really here tracking a size effect: larger countries have values of ϕ_i closer to unity. Figure 6.2 shows the relationship between ϕ_i and population in 2010 taken from the World Development Indicators database of the World Bank.

The linear fit indicates that larger countries are indeed generally more self-reliant. We can take account of this by looking at deviations from the line. These deviations represent the extent to which a country is self-sufficient (or not) *even given* the size of their population.

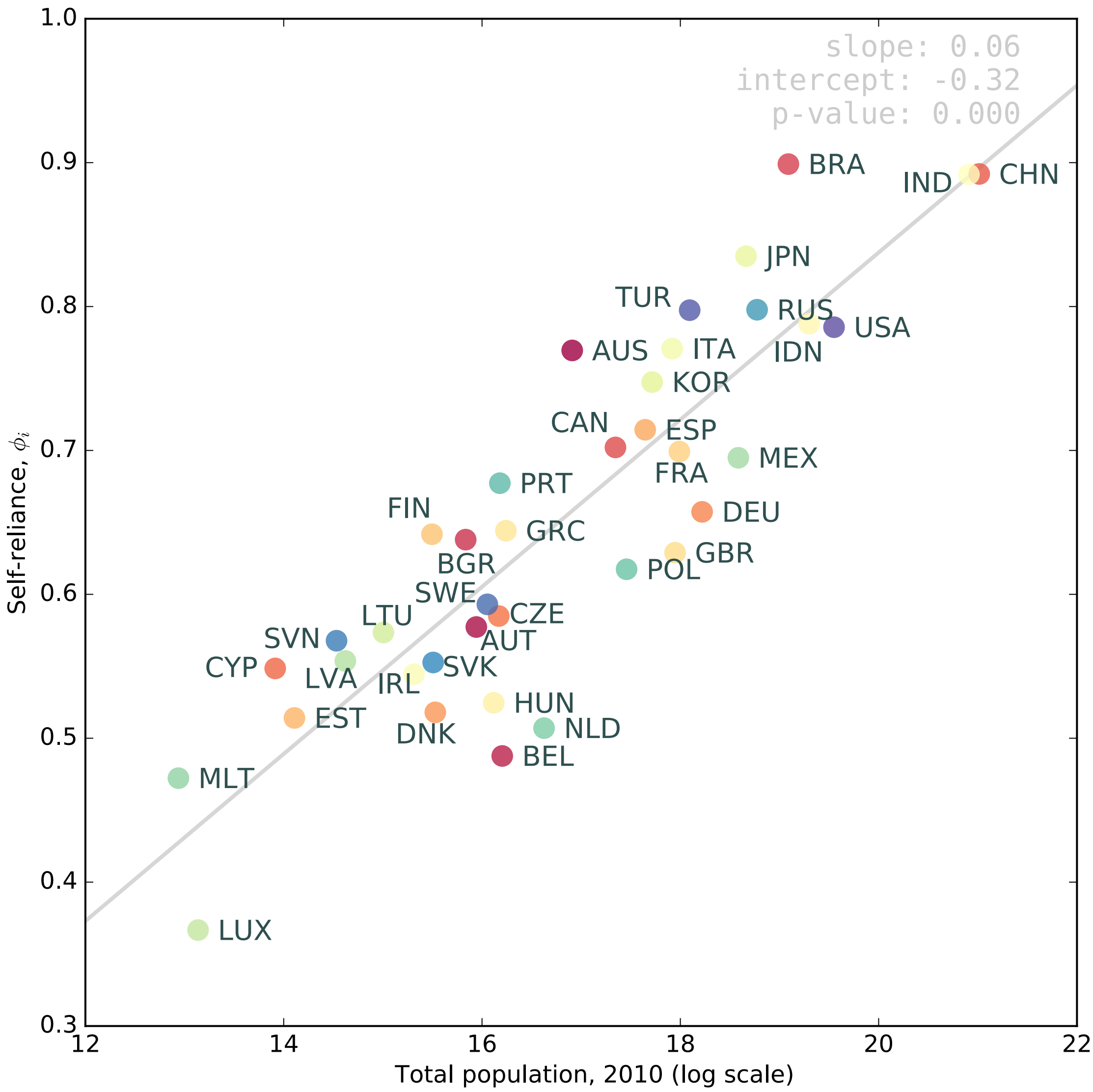


Figure 6.2: The relationship between economic self-reliance, as defined by equation (6.6), and population. Both scales are logged. The solid line shows the result of a linear fit.

The area above and to the left of the line contains more self-sufficient countries (given their size), and below and to the right contains more import-dependent. Brazil, Australia and Turkey are both more self-sufficient than their populations would suggest. Belgium and the Netherlands are less self-sufficient than expected.

Hopefully this very brief set of studies gives a sense of a kind of analysis made possible by a model of the global economy such as ours.

We have also demonstrated the first addition of extra data, in this case on populations, to bring further richness to the analyses. This idea will be greatly extended in subsequent chapters. Future research could break down the results of this section sector-by-sector or, as mentioned above, look at the changing status of these countries through time.

6.4 Import substitution

The import ratio of a CS is a measure of how much a country relies on imports of a given sector in its production and domestic final demand. It provides us with a natural way of modelling import substitution, the process of replacing imported goods with domestically produced goods in intermediate and final demands. Indeed, the 1989 *Handbook of Development Economics* gives a definition of import substitution which coincides precisely with our definition of import ratios:

> Import substitution is often "measured" by a change in the ratio of imports to the total availability (imports plus domestic output) of a single product or category of products. If this ratio falls over time, then import substitution is said to take place in that particular sector. (Bruton 1989, p. 1604)

The concept is most often associated with the study of developing countries (Bruton 1998) although it is often considered to be a failed development strategy (Baldwin 2000; Eckstein 2001; Durbarry 2004). Here though we consider it outside of the context of development. We will outline a very simple experiment, very similar to that outlined in section 6.2, but where we now decrease the import ratio of each CS in turn, rather than decreasing the final demand.

6.4.1 Decreasing the import ratio

According to equation (3.4) on p. 90, each import ratio is defined by imports, exports and total production. But these are values from data, and the import ratios are only calculated once at the time the model is created in computer memory. Once the model is created, parameters can be changed freely, without instantiation equations like this one having to hold. We are therefore free to adjust the import ratios in any way we please.

In experiments whose results are not shown here, responses to changes in import ratio were very close to (but not quite) linear, even across a -0.1% to 0.1% range[5], and therefore we will content ourselves with the results of a “small” decrease in import ratio, and not report any second-order effects.

Unlike in section 6.2 it now makes sense to use GDP as our metric to determine the effect of import substitution on the global economy. This also seems like a natural setting to divide the effects into domestic and foreign, since we would expect domestic GDP to rise and all foreign GDPs to fall. Specifically, a decrease in the import ratio of

[5] Recall that an import ratio is constrained to be between 0 and 1, and therefore a change of 0.1 is comparatively large.

Table 6.2: The 10 most significant country/sectors in terms of their effect on GDP in response to a small decrease (of 0.04) in the import ratio. Also shown are the domestic (Δ^D_{is}) and foreign (Δ^F_{is}) responses. All figures are shown in $US billions.

		$\Delta^D_{is}GDP$	$\Delta^F_{is}GDP$	$\Delta_{is}GDP$
USA	Mining	16.94	-1.88	15.06
China	Chemicals	19.42	-10.94	8.48
China	Mining	12.15	-4.39	7.76
China	Electricals	16.05	-9.14	6.91
USA	Fuel	15.57	-9.02	6.54
Canada	Financial Services	6.11	-0.12	5.99
USA	Chemicals	19.57	-14.45	5.13
USA	Agriculture	12.77	-7.65	5.12
UK	Hospitality	4.93	-0.11	4.82
Germany	Communications	4.05	0.16	4.22

country i will increase production in i, and reduce import demand. This will therefore reduce the export demand on all countries trading with i (and hence, indirectly, reduce the export demand on i itself).

As before, we will refer to changes to GDP in response to this small drop in the import ratio of sector s in country i using $\Delta_{is}GDP$, and drop the GDP. Domestic effects will be denoted with Δ^D_{is} and foreign effects with Δ^F_{is}.

6.4.2 Results

The results of this experiment are shown in table 6.2 which shows the ten largest CSs by overall response, Δ_{is}.

The USA appears most frequently in the table, with its mining sector having the largest overall response, then fuel, chemicals and agriculture. However the 2nd, 3rd and 4th spots are all occupied by Chinese sectors, namely chemicals, mining and electricals. Services sectors in Canada, the UK and Germany complete the table. There are some large discrepancies between the magnitude of the positive domestic effect and the negative foreign effect between CSs, even within a country.

The most noticeable of these is in the USA, where the country's mining and fuel sectors have similar Δ^D_{is} but hugely different Δ^F_{is}. We can see these discrepancies in more detail in the scatter plot in figure 6.3 showing the 20 largest Δ_{is}s. For visual clarity we have taken logs of the responses.[6]

[6] We of course must first take the absolute value of all negative numbers to allow the taking of logs.

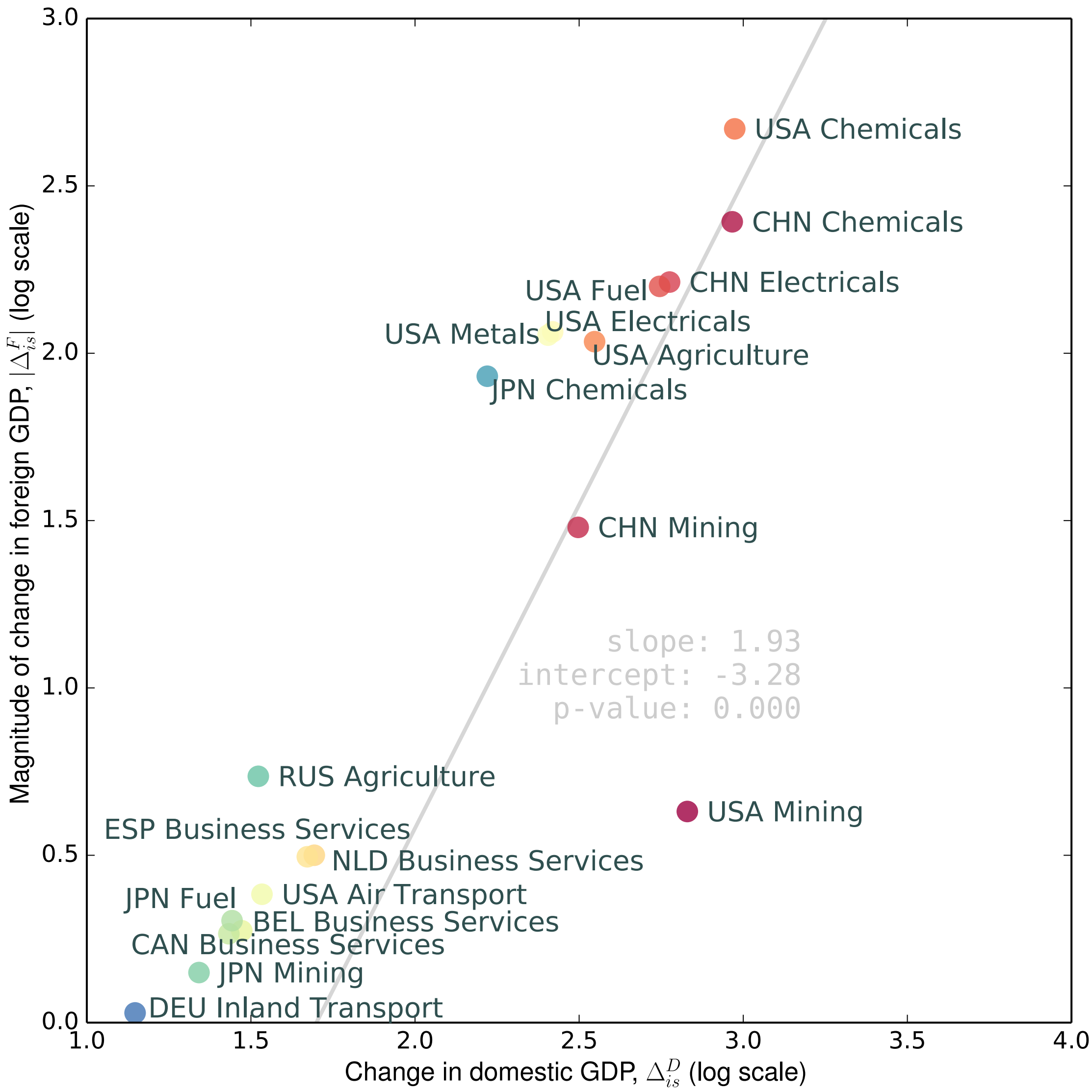


Figure 6.3: Responses to a small reduction (0.04) in the import ratio of the given CS. The responses have been divided into foreign and domestic and are shown on a log scale, with all negative numbers replaced by the equivalent magnitude to enable taking of logs. The solid line shows the result of a linear fit.

This figure shows us clearly that the US mining sector is a huge outlier, and that in fact all other CSs are remarkably close to where a line would fit without it. The area below and to the right of the line represents CSs with a disproportionately large domestic GDP increase due import substitution compared to the GDP decrease experienced by the other countries in the model. This suggests that the USA could produce more of its mining output domestically without this hurting the global economy as much as is the case with most other sectors.

Combining this approach with the extension to the 40-country version of the model used here (see chapter 5) would allow the exploration of a whole range of effects on the global economy of a particular target CS, including both the beneficial effects to the country in question and the detrimental effects to others in the target country's trade network. This idea is expanded upon at length with respect to export infrastructure in chapter 10.

6.5 Conclusion

In this chapter we have presented a number of suggestions for methods of interacting with the demonstration model to answer “what if?”-type questions of the global economy. To do this, we have manipulated some of the four sets of coefficients, namely final demand, technical coefficients, import ratios and import propensities, which are needed to solve for the complete set of trade flows, intermediate flows and metrics.

We started by reducing the level of final demand in each CS by a small amount and assessing the impact of this change on gross global output (GGO), finding that the Chinese vehicles sector is the most significant by this measure. We noted that measuring changes to GGO is not equivalent to measuring value-added-type measures such as GDP, and this is important in interpreting the results of this experiment. We found that the Chinese vehicles sector uses the most input per unit of output produced, once the entire global supply chain is taken into consideration. In fact, when we then split the effect into domestic and foreign effects, we found that China’s significance (for which we might now read *production inefficiency*) relates primarily to domestic, not imported, products.

This brought us on to a more general investigation of economic self-sufficiency, via a related measure. We found that although self-sufficiency is very closely correlated to population size, there are a couple of notable outliers: namely that Brazil is more self-sufficient than expected, and the Netherlands and Belgium are less so.

We finally demonstrated the use of import ratios in the modelling of import substitution, showing how a model of the global economy can capture not just the beneficial effects to the country in question, but also the detrimental effects to those countries whose exports suffer as a result. This is an important aspect of this kind of global modelling, and one we will return to again and again in the analyses which follow. We found that the US mining sector is a large outlier in terms of its ability to produce domestic GDP gains in response to import substitution without the typical attendant reductions in GDP abroad.

Beyond these specific experiments whose particular results suffer from the same caveats as the limited nature of the demonstration model, we hope that this chapter demonstrates the flexibility of a global model with several sets of manipulable coefficients such as ours. We have seen how manipulation of various of these sets of coefficients can allow us to begin to address policy questions, and this is a theme we will explore in the examples which follow.

Here, we have posed questions which (mostly) required no additional data to be brought to bear. In this next chapter, we present one more within-model analysis which takes an idea a little further than the brief analyses here. Then, in the remaining chapters, we will explore how adding more data can extend the scope of our investigations even further.

Chapter 7

A vector geometry approach to analysis*

7.1 Introduction

In this chapter we take a within-model approach similar to that of chapter 6, but here the analysis will be a little more involved, showing another example of the variety of policy questions we can begin to address with the coefficient sets made available by our model. Innovative approaches from the worlds of vector geometry and graph theory are introduced.

Our motivation here will be the concept of balanced economic growth, defined by Gardiner et al. (2013) as "the simultaneous and coordinated expansion of several sectors, i.e. sectoral diversification". A study of balanced growth inherently requires an approach with disaggregated sectors. Such approaches are natural in an input-output context such as ours.

We model the sectoral breakdown of each country via the relative magnitudes of each sector in that country's final demand. These relative magnitudes define a pattern which uniquely describes that country's economy in terms of demand for goods and services. These sets of final demand quantities can be viewed as high-dimensional vectors, with the demand in each sector being a "co-ordinate" or dimension of the vector. This opens up new ways of thinking about the relationship between countries.

Throughout this chapter, we introduce several innovative approaches to studying international trade in this way. Seeing the various elements of an input-output table as vector and matrices is not new. Indeed, matrix inversion is central to the technique. But here we go further into the mathematics of vectors, and vector geometry, employing

* This chapter uses the 40-country model throughout. See page 77 for a description of what this means.

techniques from those fields which are new to macroeconomic analysis. The results of this section, as with the other experiments in this part, should therefore be seen in the light of this experimental spirit: we are more interested in the application of new techniques than to the specific results we find, although we will analyse and interpret these where appropriate.

We pick a particular country to use as a base for the analysis and study changes to its final demand pattern. By viewing the set of all such sectoral breakdowns in the model as a *feasible set*, we can ask the question: is there, among the feasible set, a consumption pattern (target) which would be better for the economy of the base country than its own pattern?

By applying each proposed new consumption pattern to the base country's model, we can assess the impact on the GDP of the base country, and of all other countries in the model. This allows us to present a ranking of proposed "target" countries for each base country. Following a brief review of the literature on balanced growth, we outline in section 7.3 the process by which we "move" between base and target consumption patterns. We then show our findings in sections 7.4 to 7.6 and conclude in section 7.7.

7.2 Literature review

As Temple (2008) points out, the term balanced growth has two distinct meanings in economics. The first, not relevant to our case, refers to an equilibrium growth path in which output grows at the same rate as the capital stock. The second meaning, pertinent to the work of this chapter, relates to the idea that growth across a number of sectors is more beneficial than focusing on one sector at the expense of all others.

The discussion of balancing growth across sectors began with the discussion of UCL economist Rosenstein-Rodan (1943) about what to do with the "agrarian excess population," which at the time accounted for 25% of the population of Eastern and South-Eastern Europe. In it, he advocates a large-scale planned industrialisation across "a whole system of industries." This would allow the industry to sell its goods to workers of other industries. In doing so it would reduce the risk associated with dependence on one industry. It would also leverage what would now be called "economies of agglomeration," whereby lowered transport costs, increased competition in labour demand and accumulation of knowledge and skills can bring increasing returns to scale as the industries grow.

The concept of balanced growth is further refined by Lipton (1962) into three forms: extreme, moderate or sophisticated. The extreme view is that output in each sector (or investment, if that is the target) should "expand at equal rates in all sectors of the economy." This was broadly the view espoused by Nurske (1953). The moderate view is that sectors should all expand simultaneously though not at equal rates, an idea Lipton puts down to Lewis (1955).

Lipton says that, even by 1962, these two concepts of balanced growth were "often ignored" and that sophisticated balanced growth was the most common form.[1] This form of growth focuses on creating just enough demand to absorb the new supply of goods. As Lipton puts it, "shoemakers are neither willing nor able to buy all they produce" (p. 642). This can be achieved by using the income-elasticities of demand for the goods of each sector, and selecting growth rates for each industry whose ratios match these elasticities.

A similar argument can be made in terms of balance of supply. Since sectors require one another's outputs as intermediate inputs, it follows that simultaneous growth avoids shortages of what Lipton calls "ancillary factors."

[1] Indeed Streeten, writing three years earlier, goes further in saying saying that "balanced growth is almost axiomatic as a desirable objective, for both developed and under-developed countries" (Streeten 1959, p. 169).

These arguments are then set against the cases made for unbalanced growth by Hirschman (1958) and Streeten (1959). Hirschman's concept, according to his reviewer Shannon (1959), "is contrary to generally accepted notions of strategy but is in any case probably the only practicable approach." In it, Hirschman advocates growth through focused efforts which create "incentives and pressures for further action." He gives the example of the construction of a steel mill which would then demand of the education sector an improvement in the skills of school leavers.

Streeten describes the ubiquity of the balanced growth doctrine, but also claims to be able to hear "occasional rumblings of dissent," of which he gives an entertaining example:

> The choice between attacking on a wide front and storming key positions is a perennial one...Without knowing which or how many priority targets must be attacked to get the dynamic process of economic development under way on a cumulative basis, instinct suggests that this may be a more fruitful method than balance...But economic development in the past did not spring full blown from the brow of an economist. Nor is there any reason to think that it will be impossible to duplicate historical development which has been piecemeal (Kindleberger 1952, p. 392)

Streeten's own arguments for unbalanced growth revolve around analogies between a national economy and an individual consumer. He argues that consumption is indivisible, and therefore development must move in discrete jumps, and that there is a "growth of wants" such that consumption of one good leads to the desire for another. These effects both produce "a strong incentive for investment and further production, which would be absent had there been fully balanced growth" (p. 175). England's Industrial Revolution is called to service as an example of how lop-sided growth can lead to rapid economic progress, whereby the existence of foreign markets and foreign lending can counter Lipton's arguments for balance.

This discussion in the 1950s and 1960s led broadly to three strands of literature discussing balance between sectors. The geographical balance exemplified by much of the work of Paul Krugman, the development economics of the likes of Chenery, Robinson, and Syrquin (1986), and what de Long and Summers (1991) refer to as "the new growth theory traditions" such as Romer (1986). Romer's model is a highly theoretical attempt to explain observed increasing returns to scale, whereby large countries grow faster than small countries. The Krugman geographic sense of balance is not of focus here, so we present a short overview of the development strand beginning with Chenery.

Chenery, Robinson, and Syrquin (1986) look at how important industrialisation is for development, the relative importance of supply-side vs. demand-side changes in producing growth, and what the causal links between these phenomena might be. To this end, they examine trade theory and how it might affect external policy making, and internal policy-making and resource allocation via the study of multi-sector models, both input–output and CGE. It is the second of these which are of relevance to us here. The authors argue that this approach "is particularly suited to the study of industrialization because the indirect effects of structural change often outweigh the direct effects that are visible in the national accounts aggregates." (p. 6).

They present a "multisector simulation" of structural transformation (Chenery and Syrquin 1986) and an input-output model of changes in the pattern of demand, trade and production (Kubo et al. 1986). The first of these is an input–output study comparing nine countries, each with 23 sectors.[2] The accounting framework they use is standard but the relationships between the elements in the input–output table goes beyond the standard Leontief-production/matrix-inversion approach.

Each element of gross national product (consumption, investment, government spending, exports and imports) is expressed as a function of "universal" factors, where every country is assumed to behave in the same way, and "particular" factors which distinguish countries from one another. The universal factors are per-capita income, capital stock and "skills." The particular factors are population, N, natural resources per capita, R, and "allocation policies," denoted Φ. Domestic demand is modelled following the linear expenditure system (see section 2.7.4 on p. 54). To model the process of technology and factor substitution as development occurs, the input–output coefficients themselves are modelled as being functions of per-capita income, as is total value added.[3]

In each case, both the "universal" and the "particular" parts of each functional forms is estimated separately in a panel regression. The form is of a second-order approximation using, for the example of the particular parts, the following formula where x is the variable of interest:

$$x = \alpha + \beta_1 \ln y + \beta_2 (\ln y)^2 + \gamma_1 \ln N + \gamma_2 (\ln N)^2 + \varepsilon F \qquad \text{(eq. 3-14, p. 46)}$$

where F is the trade balance, exports minus imports.

---

[2] The 23 sectors are aggregated into eight sectors for presentational purposes.

[3] The example the authors give is of agricultural requiring increasing intermediate inputs as mechanization replaces "hand operations" (p. 45).

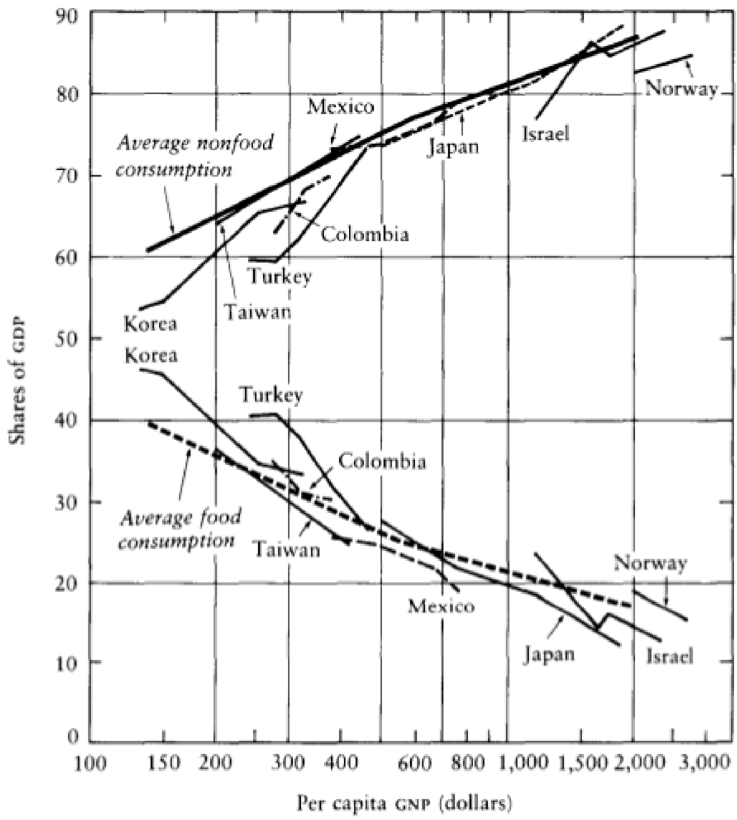


Figure 7.1: Figure 3-2 from Chenery and Syrquin's multisector simulation of the pattern of growth in nine countries, tracing the share of food and non-food sectors in GNP as countries develop. The average change is shown in black.

The process of industrialisation is modelled as a sequence of increasing levels of income per capita with six discrete levels (see the authors' Table 3-3 on p. 48). In this way, a forecast growth path for each of the nine countries can be traced out and compared country by country.

By way of a final note on this book chapter, the authors include an interesting method of capturing the results of a model with many outputs. For example, figure 7.1 shows the output of the simulation for each of the nine countries, showing how food and non-food shares of GDP change for each of the nine countries, along with an average change. This style of visualisation is extremely impressive for its time and appears to have been drawn at least partly by hand.

The other chapter of Chenery, Robinson, and Syrquin (1986) of relevance here is that of Kubo et al. (1986). Here exogenous changes in demand, trade and production are analysed in terms of their effect on the sectoral composition of aggregate output. The focus of study are the effects on this of the input–output technical coefficients, what Kubo calls “import coefficients” and the composition of trade and final demand.[4] The change in output of sector i is modelled as the sum of five exogenous structural changes: an increase in domestic demand; an increase in exports; an decreased import coefficient on final goods, and on intermediate goods, both implying greater use of domestic products; and changes in input–output coefficients. Each of these exogenous changes is “fed through” the input–output table to find a proportional response of total output.[5]

This change-decomposition approach is then combined with the model results of Chenery and Syrquin (1986) outlined above. Each countries simulated growth path is analysed in terms of the deviation of changes in sectoral output from what would be expected under a balanced (i.e. proportional) growth scenario. This allows the authors to decompose the sources of sectoral deviation for each of the nine countries. This gives each country a unique “fingerprint” showing how each of the five structural changes contribute to the deviation from balanced growth simulated by the model.

The work of this book went on to inspire a large literature on the nature of industrialisation, “vertical specialisation”–the extent to which a country specialises in one part of a longer production process, and the role of technology and diversification in development more generally. Of particular note in this context are the following: Imbs and Wacziarg (2003) who use a similar framework to show that sectoral diversification first increases as countries develop but then decreases beyond a certain level of income per capita, a finding more recently supported by Clark and Sawyer (2014); Davis (1995) who looks for evidence fo the so-called “Dutch disease,” whereby a sudden increase in the growth in one sector, in Davis’s case minerals, can actually harm an economy rather than prompt growth across other sectors. His approach is purely statistical, with no modelled theory, but he nevertheless comes down strongly in favour of developing countries indeed exploiting their natural resource wealth, suggesting that the Dutch disease has not affected the countries in his study; and Kildegaard and Williams (2002) who equate industrial concentration with systemic risk, the idea being that a non-diversified economy will cause the negative effects of productivity or demand shocks in a sector to be unmitigated by other, unaffected sectors.

All of the reviewed literature in this review implicitly assumes that a theoretically desirable sector mix is also achievable in reality. In this chapter, we take a different approach, assuming that only sectoral patterns which exist in the world are feasible. In

4 As outlined on p. 123 of Chenery, Robinson, and Syrquin (1986), there are two sets of import coefficients. There is a single “domestic supply ratio” for intermediate use, u_i^w, and a separate ratio, u_i^f, for final use. The subscripts refer to the supplying of sector i’s goods, and the ratios are assumed to be the same across the demanding sectors. The surprising w superscript comes from the fact that, in the chapter which introduces these assumptions, Chapter 5, W is used for intermediate use.

5 A much simpler version of this analysis is carried out in section 6.2.

other words, the complete set of existing sectoral patterns represents a “feasible set” from which countries seeking to adjust their sectoral mix can select.

7.3 Perturbing final demand vectors

The mathematical ideas of this section, and some of the language used to express them, are due to work conducted in collaboration with Robert J. Downes. This is particularly true of section 7.A, the mathematical appendix.

Each country i in our model has a level of final demand for each sector s which we have denoted as f_{is}, which describes how much (in dollars) that country's population and government consumed of the output of that sector in the year upon which our model has been based.[6] By stacking these f_{is} vertically for each sector, we arrive at a country-specific final demand vector with S elements, or *dimensions*, where S is the number of sectors. We call this an S-vector, and it is represented as follows:

$$\mathbf{f}_i = \begin{bmatrix} f_{i1} \\ \vdots \\ f_{iS} \end{bmatrix} \tag{7.1}$$

The set of all 40 final demand vectors forms the *feasible set* of consumption patterns. As with all vectors, this vector has a length and a direction. We use these geometric ideas to compare countries' final demand vectors with one another.

Initially we specify a *base* country whose final demand vector we will analyse. Then, every remaining country in the feasible set becomes a *target* country: we rotate the base country's final demand vector in the direction of each target country in turn, ensuring that total consumption in the base country, $\sum_s (f_{is})$, remains fixed. In effect, the elements of the final consumption vector are adjusted up and down in such a way that the sum of consumption levels is constrained to stay fixed. Using the language of vector mathematics, we refer to these adjustments as *perturbations*.

We evaluate the impact of the perturbation on the base country's GDP by running our global model using the perturbed consumption pattern. For every base-target country pair we then record the magnitude of the change in base-country GDP produced by perturbing the base country's consumption pattern towards that of the target country; this may be positive or negative. We can repeat this process taking every country in turn as the base, and comparing it to every target creating a full set of base/target pairs.

6 Here, and throughout, the 40-country WIOD model is based on data from 2010, unless otherwise specified.

As a demonstration of this process, a sample set of model runs is given in figure 7.2. The base country is Great Britain (GBR) and the % change in GDP resulting from each perturbation is recorded for each target country in the feasible set as the vector rotates from its base to its target angle. The horizontal axis may be thought of as the *distance* the base-country vector has moved towards the target vector; units are explained in section 7.A. Thus, each line in figure 7.2 traces out the change in GDP in GBR in response to a movement of GBR's final demand vector towards that of the target country labelling the line. Each line stops when GBR's final demand vector is colinear with, i.e. pointing in the same direction as, the vector in the target country. In economic terms, when GBR consumes in precisely the same pattern as in the target country. What is crucial about this figure is that it shows that changes to GDP are monotonic (though not linear) with respect to the rotation angle. This means that a study of small rotations will be sufficient to get a sense of the picture across all larger rotations.

Perturbations towards Great Britain's nearest target country, Germany (DEU), lead to a decrease in GDP; if Great Britain consumes exactly like Germany, the magnitude of this decrease is approximately 3%. Conversely, consuming exactly like Denmark (DNK) increases Great Britain's GDP by over 2%.

When running the model in what follows, we will standardise the perturbation distance to 1 deg throughout our analysis;[7] this provides a measure of comparison across all countries.

For this fixed distance, we rank target countries by the number of base countries they improve, i.e. the number of perturbations towards each target country resulting in a positive change to the base country's GDP. We also rank base countries by the number of target countries by which they are improved.

In the next section we take the innovative step of treating these rankings among countries as a *network*: countries are improved by, and improve, other countries. This forms a network, or graph, of relationships, allowing us to bring some of the tools of graph theory to the study of global trade.

[7] In radians, the units of figure 7.2, 1 deg $\equiv \pi/180$ rad.

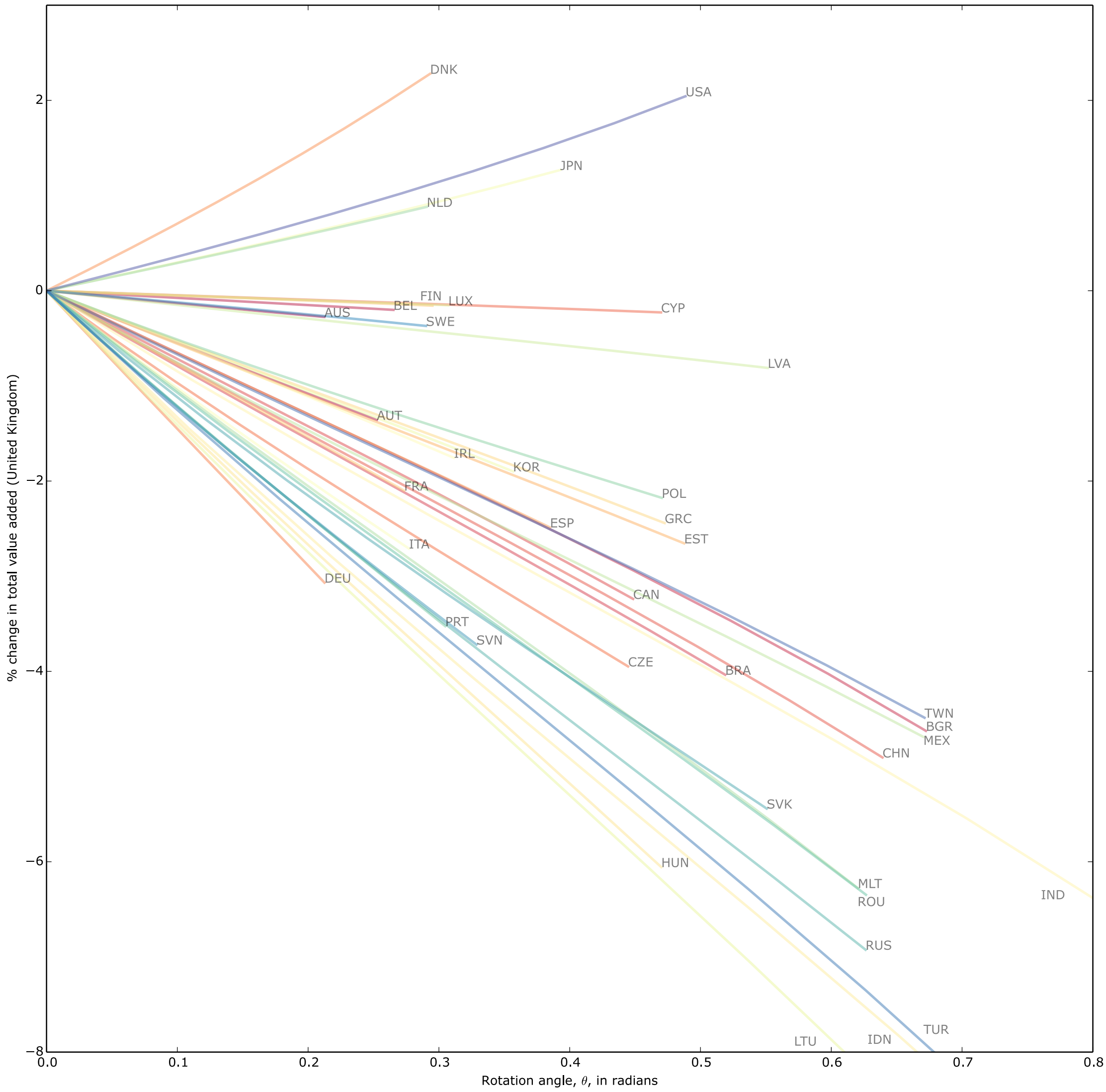


Figure 7.2: Perturbation distance (rotation angle) against % change in GDP for Great Britain. All trajectories originate at Great Britain's consumption pattern ($\theta = 0$). Destinations are the consumption patterns of each target country in the feasible set labelled by ISO alpha-3 code.

7.4 Analysis: directed network

For each base-target country pair, the perturbation process detailed in section 7.3 produces a perturbed consumption pattern, without changing total consumption. The result of evaluating this perturbed consumption pattern using our global model is a modified GDP figure, the outcome of the base country perturbing its consumption pattern towards the target country's consumption pattern. For brevity, we refer to this process as the base country consuming *like* the target country.

If the GDP of country i increases when it consumes like country j we say that "i is improved by j". Then, to each country i is associated a set of countries by which it is improved and a set of countries by which it is *not* improved. We can arrange these relationships in a matrix G, with i and j indexing rows and columns, respectively:

$$G = \begin{cases} 1 & \text{if } j \text{ improves } i \\ 0 & \text{Otherwise} \end{cases} \tag{7.2}$$

This matrix G can then be thought of as the adjacency matrix of a *network of improvements*: countries are the nodes of this network, and an edge exists between i and j if j improves i.[8]

As we are working with a directed network, there are two basic quantities associated with each node: the *in-degree* and the *out-degree*, which have the following interpretation: the in-degree, k_i^{in}, is the total number of countries that are improved by consuming like country i; the out-degree k_j^{out} is the total number of countries that improve country j. Table 7.1 shows countries ranked by in- and out-degree, respectively.

We see from these results that countries which are good *improvers* (high in-degree) tend to be improved by few (low out-degree), and vice versa, but that the correlation is not perfect. Figure 7.3 shows this relationship explicitly: there is a clear negative correlation between in- and out-degree metrics.

In addition, we observe by simple visual inspection that countries with high in-degree also tend to be richer than those with low in-degree. To check whether in-degree merely reflects a country's wealth, we next perform a simple statistical analysis of the determinants of in-degree.

[8] The key concepts from network theory utilised in this chapter are given in Newman (2010); we direct the interested reader to this text for an accessible introduction to the subject.

Table 7.1: Results of consumption pattern perturbation for all base-target country pairs. In (a), target countries are ranked by the number of base countries they improve (in-degree). Conversely, (b) shows base countries ranked by the number of improving target countries they have (out-degree).

(a) Target countries ranked by in-degree

Country	In-degree
USA	36
Denmark	36
Japan	35
Netherlands	30
Great Britain	27
⋮	⋮
Hungary	3
Russia	3
Lithuania	2
Indonesia	1
Turkey	1

(b) Base countries ranked by out-degree

Country	Out-degree
Turkey	39
Lithuania	38
Russia	37
Hungary	32
Malta	31
⋮	⋮
Luxembourg	2
China	1
Indonesia	1
Japan	1
USA	1

7.4.1 Sectoral diversity

As outlined in this chapter's literature review, the central thesis of balanced economic growth is that a more diverse economy leads to stronger economic growth (Murphy, Shleifer, and Vishny 1989). Under this assumption we would expect a country which improves many others to have a diverse final demand vector, in other words, one not dominated by a small number of sectors.

We use the standardised M1 index of qualitative variation (Gibbs and Poston 1975) which, gives the likelihood of any two samples from a population fall in the same category. In this context, it measures the likelihood of two randomly selected dollars of final demand expenditure occur in the same sector. In effect, it expresses the tendency of an economy to be dominated by a small number of large sectors. It is from the same family of measures more commonly known to economics via the Gini index of wealth inequality or via ethnolinguistic fractionalisation, a measure of how divided a country is along ethnic and linguistic lines. More specifically, the M1 index was used in the Simpson diversity index (Simpson 1949) and the Herfindalh-Hirshman index (Hirschman 1964), commonly used to identify company market share within a given sector.

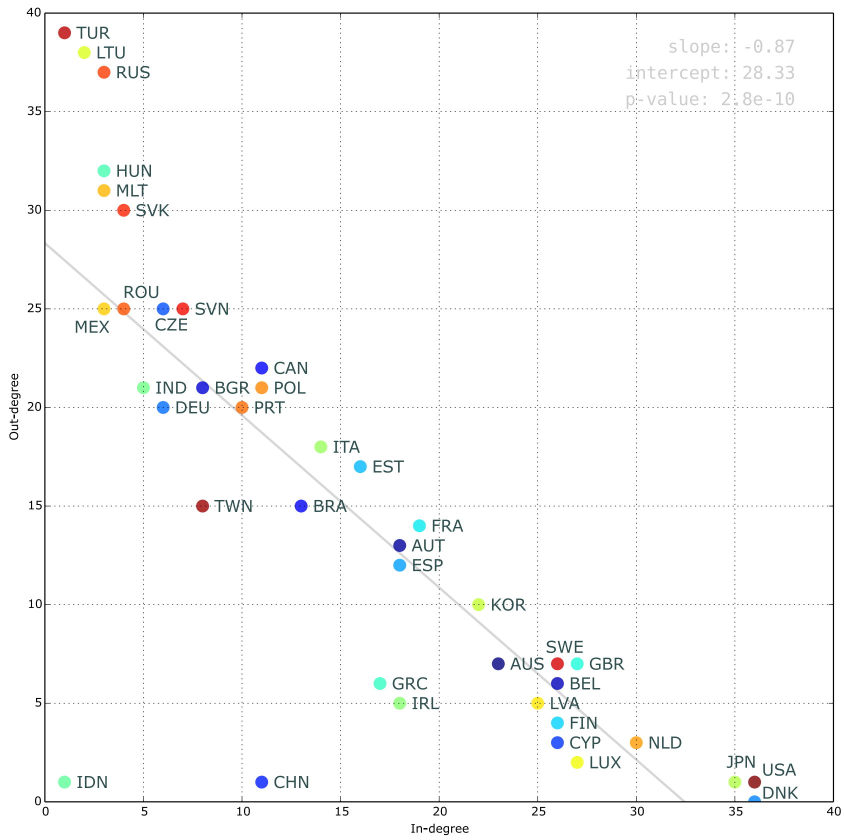


Figure 7.3: In-degree, a count of the number of countries who would benefit from consuming like the country in question, plotted against out-degree, a count of the number of countries the country in question would benefit from consuming like. A linear fit has been added to the plot.

The measure is defined as:

$$\mathrm{M1}_i \equiv 1 - \frac{\sum_s \left(p_{is}^2\right) - \frac{1}{35}}{1 - \frac{1}{35}} \tag{7.3}$$

where p_{is} is the share of final demand held by sector s of country i when the economy is partitioned into 35 sectors; for each country, p_{is} is determined directly from the corresponding final demand vector as:

$$p_s = \frac{f_{is}}{\sum_r f_{ir}} \tag{7.4}$$

Note that $\mathrm{M1} = 1$ when the share of the national economy held by each sector is uniform; qualitatively, this is a perfectly balanced economy. Conversely, $\mathrm{M1} = 0$ when the economy is dominated entirely by a single sector; qualitatively, this is a completely unbalanced economy.

7.4.2 Determinants of a balanced economy

The relationship between sectoral diversity and in- and out-degree, respectively, is investigated in figure 7.4. According to the balanced growth literature, we might expect that a balanced economy acts as a good improver. We do indeed see a positive relationship between M1 and in-degree. The most improving countries from table 7.1, the USA, Denmark, Japan and the Netherlands, are also the most balanced. There is also a cluster of unbalanced countries which are very poor improvers: Russia, Slovakia (SVK), Slovenia (SVN), Hungary and Czech Republic. Turkey and Lithuania (LTU) are slightly better balanced than their improver status would suggest, although they are still towards the lower end of the balance scale.

The most balanced countries tend also to be richest in the model. The relationship shown in figure 7.4 between M1 and in-degree may therefore simply be proxying for an underlying dependence on GDP per capita.

To allow for this possibility, we control for country-specific relationships by running a simple linear regression of the form:

$$\text{IN-DEGREE}_i = \alpha + \beta \text{M1}_i + \gamma_i \mathbf{x}_i + \varepsilon_i \tag{7.5}$$

where M1 is the measure of sectoral diversity from equation (7.3) and $\mathbf{x}_i$ is a vector of country-specific variables.

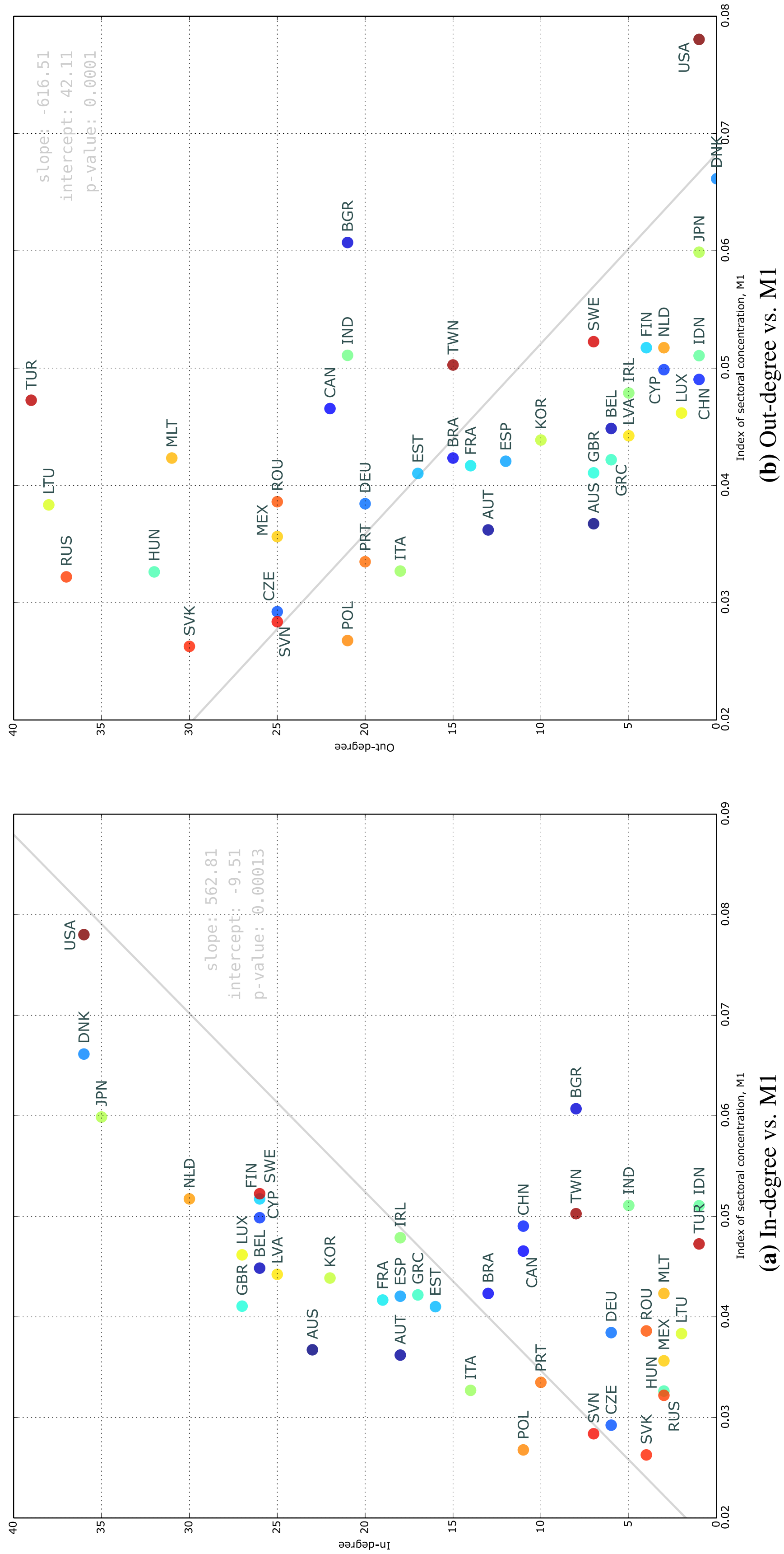
The thesis that a balanced economy makes a country more likely to be an improver implies $\beta > 0$, which should be robust to the addition of arbitrary elements of $\mathbf{x}_i$.

We give three different regression model specifications:

1. includes M1 only as the independent variable;
2. includes, in addition to M1, GDP per capita and population;[9]
3. includes, in addition to the three regressors of specification 2, two indices describing the bias of country i's final demand vector in favour of services and manufacturing sectors. These indices are calculated as the ratio of, respectively, services and manufacturing final demand to total final demand in country i. See table 7.4, at the end of this chapter, for our categorisation of the 35 economic sectors as services or manufacturing (or neither).

The summary of this analysis, presented in table 7.2, identifies the M1 sectoral diversity measure as significant in all three specifications. This goes some way towards support-

[9] The order of magnitude multiplier against each independent variable is chosen simply to result in point estimates of convenient magnitude.



(b) Out-degree vs. M1

(a) In-degree vs. M1

Figure 7.4: Figure (a) shows in-degree plotted against the M1 metric of sectoral diversity; Figure (b) shows out-degree plotted against the M1 metric of sectoral diversity. A linear fit has been added to each plot.

Table 7.2: Linear fits. Dependent variable is IN-DEGREE$_i$ in all specifications. OLS standard errors are shown in brackets below each point estimate.

	(1)	(2)	(3)
Sectoral diversity, M1 (%)	5.6***	4.7***	2.6***
	(1.3)	(1.1)	(0.59)
GDP per capita (10,000s of $)		2.4***	0.63*
		(0.56)	(0.36)
Population (billions)		−3.5	0.84
		(4.3)	(2.2)
Services Index (%)			0.76***
			(0.13)
Manufacturing Index (%)			−0.98***
			(0.28)
R^2	0.32	0.62	0.91
Observations	39	39	39
Note:	*p<0.1; **p<0.05; ***p<0.01		

ing the thesis that balanced consumption is an important criterion for an economy to be an improver.

GDP per capita is significant in specification 2, while population is not. It seems that size alone is not sufficient to make a country a good improver: that country must be rich in terms of GDP per capita.

In specification 3, both economic indicators are highly significant, but with opposing correlation coefficients: being services-biased increases improver status, while being manufacturing-biased decreases it. The addition of these indices greatly reduces the point estimate against GDP per capita, suggesting that what may appear to be a wealth effect is in reality a *bias-towards-services* effect. The services industries tend to have a relatively high per-unit value added. It makes intuitive economic sense that consuming more services and fewer manufactures would benefit a value-added measure like GDP.

Our next piece of analysis extends the network representation of this section which was a simple "does this country improve another or not?" to include a measure of the extent to which a target country improves a base.

Table 7.3: Eigenvector centrality by country for the network defined in section 7.5. Only the ten largest are shown.

Country	Eigenvector centrality
Denmark	0.960
USA	0.273
Japan	0.049
Netherlands	0.040
Luxembourg	0.011
Belgium	0.007
Cyprus	0.005
UK	0.005
Finland	0.004
Latvia	0.004

7.5 Analysis: weighted directed network

We can extend the directed network representation of section 7.4 by considering not only whether one country improves another, but also the extent to which it does so.

Formally, denote by g_{ij} the percentage change in base country i's GDP resulting from a perturbation towards target country j. As before, we arrange the g_{ij} as a matrix G', setting negative entries to zero:

$$G' = \begin{cases} g_{ij} & \text{if } j \text{ improves } i \\ 0 & \text{Otherwise} \end{cases} \tag{7.6}$$

G' may then be interpreted as the adjacency matrix of a corresponding weighted directed graph, the network of percentage GDP increases resulting from the perturbation process.[10] The weighted directed network of improvements is given in figure 7.5. In this representation, edge weight is proportional to g_{ij}. Node size is proportional, by area, to population, and the colour is related to the GDP per capita on a scale where richer countries are blue, and poorer countries are red.

By visual inspection, connectivity in this graph appears to be dependent upon GDP per capita; countries with high GDP per capita (blue-green) appear to be clustered on the left of figure 7.5, while lower GDP per capita countries (red-orange) appear on the right. We can verify this, by applying formal network clustering techniques.

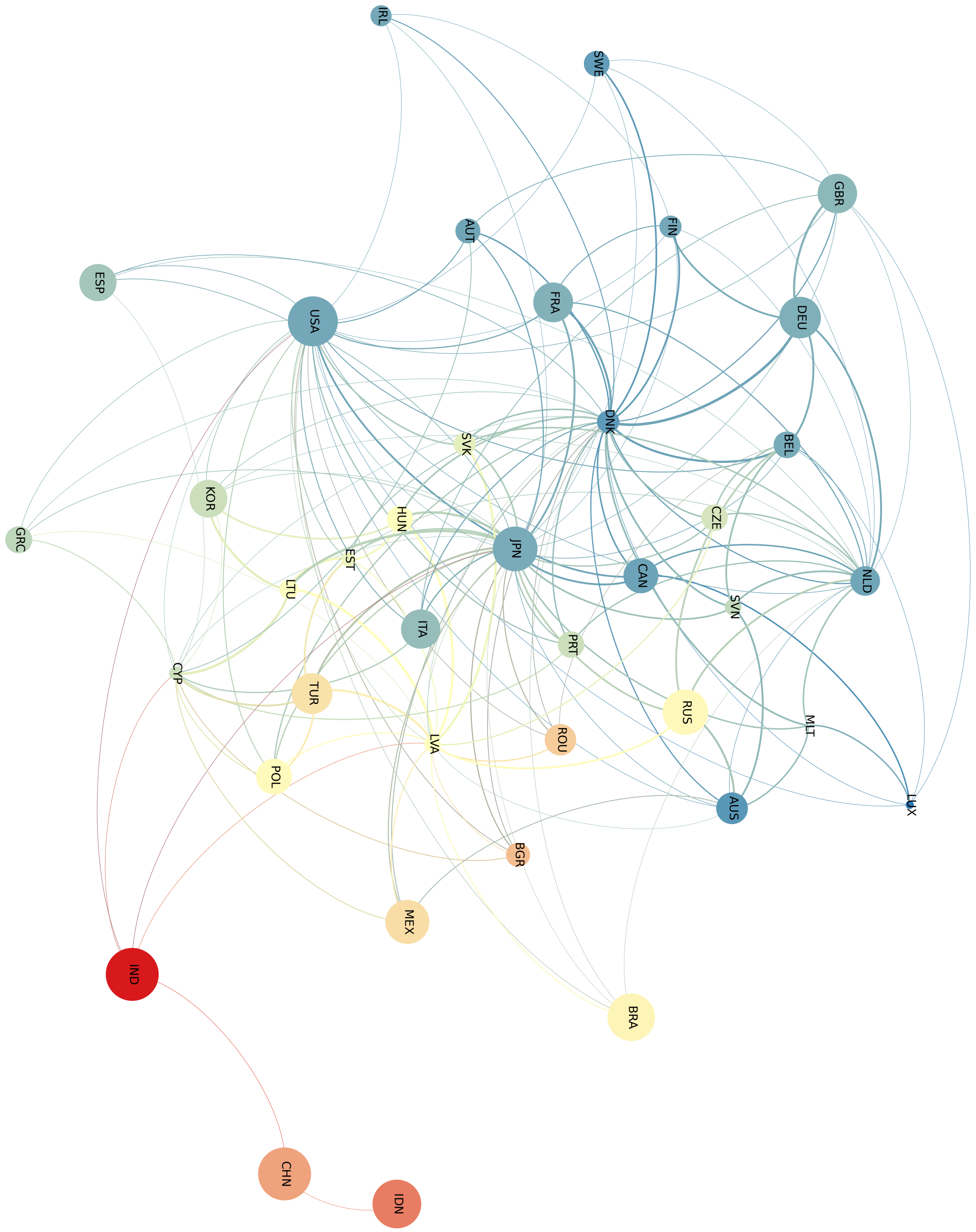


Figure 7.5: A network of improvements with countries as nodes and edges defined by equation (7.6). Outgoing links lie anti-clockwise of the direct line between two nodes (only the five largest are shown). Node size is proportional to population; colour is defined by GDP per capita set along a scale where red and blue indicate lowest and highest, respectively.

7.5.1 Centrality and in-degree

Beyond the visual inspection of the previous section, we can use the tools of graph theory to investigate the structure of the network of improvements in a quantitative manner. To begin, we consider the *eigenvector centrality* of the nodes in our network.

Generally speaking, centrality is a measure of the relative influence of a node in a network in terms of its connections to other nodes (Newman 2010). Eigenvector centrality is a refined version of this concept: the measure determines whether a node is strongly connected to the parts of the underlying network which are themselves strongly interconnected.

In our network, a country with high eigenvector centrality improves countries which themselves improve many others. Low eigenvector centrality implies that a country not only improves few other countries, but that those few improve few countries themselves. The centralities associated with the nodes in our improvers network are presented in table 7.3.

It is interesting to comparing table 7.3 with table 7.1a, where we listed the most-improving countries by a simple binary measure: do they improve a particular target country or not? We see an overlap between the two tables: the top four places are occupied by Denmark, the USA, Japan and the Netherlands in both cases. But this second way of measuring improvers offers a higher fidelity perspective. For example, we are able to distinguish between the relative influence of the USA and Denmark which, in the unweighted case, table 7.1a, were scored identically. In contrast, in table 7.3 Denmark achieves a significantly higher centrality score than the USA. Luxembourg also shows itself to be a strong improver here, stronger than the UK, although it does not appear in the top five of table 7.1a.

Our final piece of analysis takes further the idea alluded to earlier that countries appear to form clusters in the network representation shown in figure 7.5.

[10] If any. Notice that if a target country makes a source country's GDP *worse*, this effect is not captured by G' beyond the zero entry.

7.6 Analysis: community detection

While centrality measures provide information related to the relative influence of nodes within a network, they do not provide detailed structural information about the network as a whole. In order to investigate this structure further we utilise cluster analysis, which seeks to identify groups of nodes which are strongly connected to one another *and* weakly connected to the rest of the network.

Cluster analysis allows us to locate each country within a specific group, the nodes of which share stronger relationships than with the rest of the network. We then investigate the structure of each cluster in detail to determine whether the constituent countries share common economic attributes. In particular, the measure of sectoral diversity and indices given in section 7.4.1 form the basis for this analysis.

7.6.1 Method overview

We follow the method outlined in Reichardt and Bornholdt (2006a) whereby the minimisation of a certain *quality function* produces a nodal clustering which minimises the overall energy of the system. This is shown to be equivalent to the more usual modularity method presented in (Newman and Girvan 2004b); we have selected (Reichardt and Bornholdt 2006a) for the purposes of computational efficiency.

According to (Reichardt and Bornholdt 2006a), the quality function should "follow the simple principle: group together what is linked, keep apart what is not". Assessment of the quality of a given clustering of nodes is carried out relative to an underlying null model, determined from the in- and out- degree of each node in the network.

The relationship between the actual network and the null model is mediated by a single parameter, γ. There is no guidance given in the original work for selecting an appropriate value for γ. We therefore select an appropriate γ value using a parameter-sweep to find the value most likely to produce precisely two clusters of countries. This is to simplify the analysis as far as possible. The results of this parameter sweep are shown in figure 7.6.

Because the algorithm, as with all clustering algorithms, is highly stochastic, 50 runs were performed at each level of γ. After each run, the number of countries in each cluster was recorded. For each of the 50 sets of runs at each value of γ, the mean and median number of countries per cluster was calculated. Corresponding to two equally-sized clusters, we would want a median number of countries per cluster of 20. The figure shows that the algorithm is clearly sensitive to this parameter, with low values (below around 0.2) resulting in the median group size fluctuating between 40

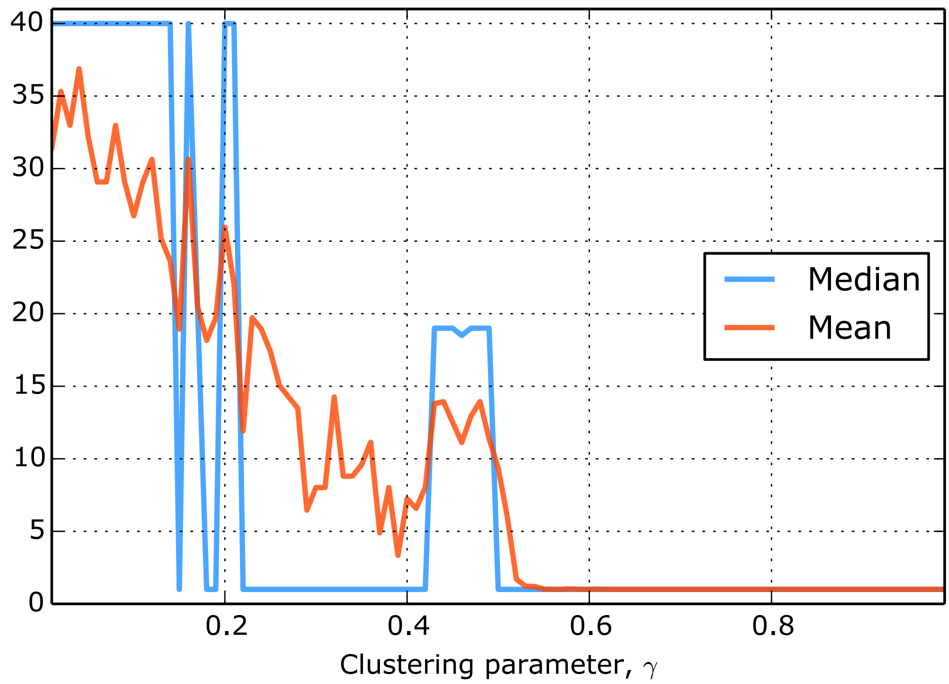


Figure 7.6: Results of multiple runs of the clustering algorithm of Reichardt and Bornholdt (2006a) with various values of the parameter γ. The algorithm was run 50 times at each level of γ and the mean and median number of countries per cluster was recorded. The results show a clear spike around$\gamma = 0.45$.

and zero. This region is of no interest to us, since all countries are being placed in the same cluster. But there is an interesting spike (robust to several runs of the overall experiment) around $\gamma = 0.45$ where the median group size is just under 20. This is commensurate with there being two approximately equally-sized clusters. We therefore use the value 0.45 in the analysis which follows.

Having decided on our parameter values, we can now run our clustering algorithm. Due to the stochastic nature of any clustering algorithm, it is prudent to run it many times. By tracking the cluster membership of each clustering run, we can record each time a pair of countries are sorted into the same cluster. This results in a matrix representation of multiple clustering runs with the ij^{th} entry representing the frequency of countries i and j being in the same cluster at the end of the process. This is called a co-occurrence matrix.

We order this co-occurrence matrix using the same seriation algorithm described in section 3.5.1 to allow us to visually inspect of the clustering of our network. This ordered co-occurrence matrix is shown in figure 7.7, and is analysed in detail in the next section.

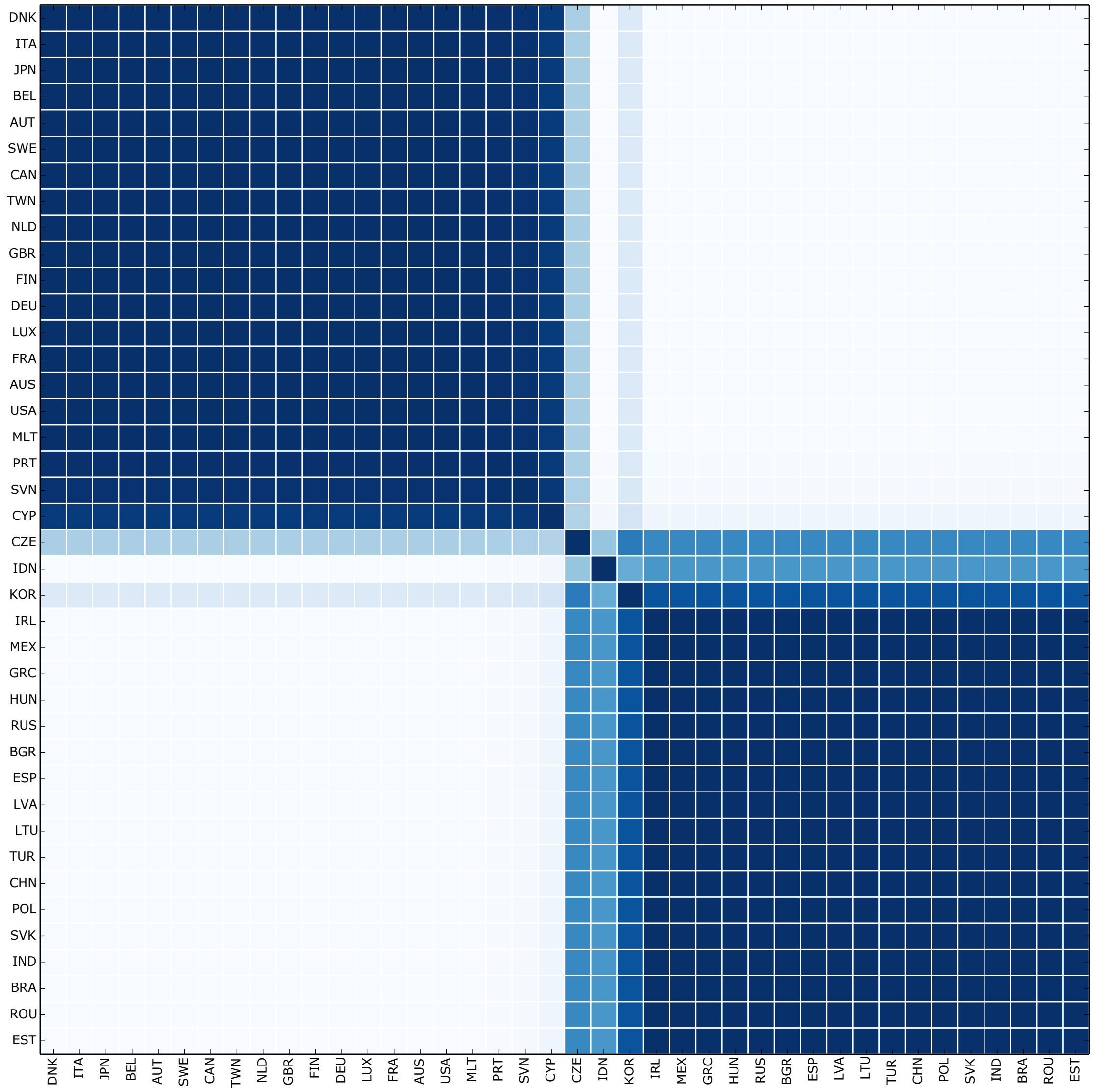


Figure 7.7: Co-occurrence matrix of the optimal clustering. The rows and columns have been ordered to maximise the similarity between them, using a process called seriation. White areas are very dissimilar; dark areas are highly similar. We see two well-defined clusters, along with a small number of countries without clear cluster membership.

7.6.2 Identifying clusters

Two very clear clusters appear in the ordered co-occurrence matrix. One in the upper-left and one in the lower-right portions of the matrix. Visual inspection suggests there may be GDP per capita disparity between the two clusters, the uppermost, with countries such as Denmark (DNK) and Germany (DEU), having high GDP per capita relative to the lower cluster, which has countries has such as Mexico, Turkey and India.

We can make this more precise by running a regression of cluster membership, a binary variable, against GDP per capita. A logit regression (results not shown) confirms that GDP per capita is indeed a strong predictor of cluster membership. This finding is robust to the use of both specifications (2) and (3) of table 7.2, our regression analysis in section 7.4.2 where we were looking for the determinants of a balanced economy. We refer to these clusters as "rich" and "poor," for brevity.

Additionally to the two main clusters, we can to identify several countries without clear cluster membership: Korea (KOR), Indonesia (IDN), and the Czech Republic (CZE) are the most clear-cut examples although further analysis suggests there are actually eight countries without a 'perfect' affinity for either cluster.[11] We might label these not as either "rich" or "poor" but as *transition* countries which fit neatly into neither category.

The role of these three transition countries can be discerned when we explicit draw out the network as a collection of nodes and edges. If we remove all edges which running between nodes of different clusters other than the transition countries, the result is as shown in figure 7.8. As in figure 7.5, the nodes are joined to those countries they improve by edges whose weight is proportional to the extent of the improvement.

We clearly see the rich cluster on the left and the poor cluster on the right, with only Ireland (IRL), Spain (ESP) and possibly Greece (GRC) being in the "wrong" cluster.

Our transition countries, Czech Republic (CZE) and Korea (KOR) are connected to both the low and high GDP per capita clusters. Close inspection of the arrow heads at the ends of the edges connecting these nodes suggests they may play different roles. Korea (KOR) appears to have mainly incoming egdes: this suggests it acts as an improver for countries belonging to both the low and high GDP per capita clusters.

[11] These countries are: Ireland (IRL), Korea (KOR), Indonesia (IDN), Czech Republic (CZE), Cyprus (CYP), Slovenia (SVN), Portugal (PRT), and Malta (MLT). Note that our decision to choose the γ parameter such that as close to two balanced clusters were produced as possible may play a role in precisely which these countries turn out to be.

By comparison, the Czech Republic (CZE) improves countries in the low GDP cluster but is improved by countries in the high GDP cluster. In this sense, it acts as a bridge between the two clusters.

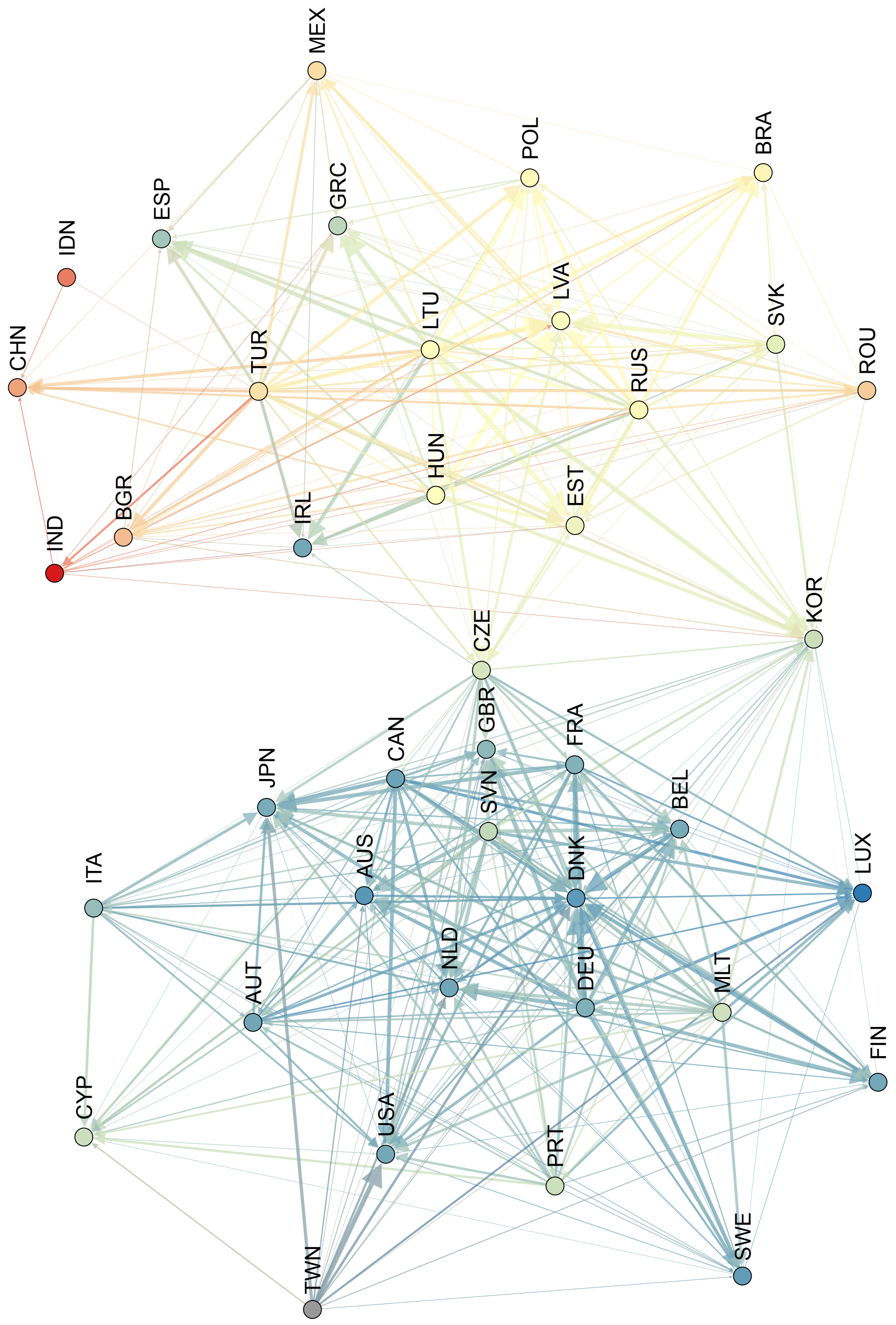


Figure 7.8: A network of improvements showing the clusters of figure 7.7. Edges between clusters have been removed, with the exception of three transition countries Korea (KOR), Indonesia (IDN), and the Czech Republic (CZE). Only improvements producing changes greater than 0.1% of GDP are shown. Node colouring reflects underlying GDP per capita; edge colouring reflects averaged GDP per capita of base and target country.

7.7 Conclusion

The work of this chapter is designed to show that by introducing innovative techniques, such as those from vector geometry and graph theory, a deeper analysis than those of chapter 6 is possible even just working within the model, with the addition of no external data sources.

We have demonstrated how a model of global trade can naturally be viewed as a network, and that the tools of graph theory can reveal interesting underlying structures when the results of a given model run are too large to be otherwise presented in a concise way.

We began by showing how a set of sector-level final demand values can be thought of as a vector, the length and "direction" of which uniquely identifies a country's consumption pattern among all the countries in the model.

We outlined a mathematical method for rotating one final demand vector in the direction of another, and showed how this causes a change in the GDP of the rotated country. We also showed graphically that this change is monotonic (see figure 7.2). We then took each country in turn as a "source", and rotated its final demand vector 1 degree towards each other country, or "target".

On the first run of this experiment, we recorded whether or not a target country increased the GDP of a source country, and labelled such targets as improvers. This allowed us to rank countries in terms of the number of countries they improved, and showed that good improvers were improved by few other countries. We also showed that having a diverse final demand vector, one not dominated by a small number of sectors, was correlated with being an improver.

This was investigated further using a statistical analysis. We showed that the sectoral diversity measure was highly significant in explaining a country's improver status, as was being a service-led economy. Being a manufacturing-led economy had a negative effect of improver status, and GDP per capita was only weakly correlated.

We then ran the experiment again, this time gathering the extent to which a target country improved a source country, giving us a more subtle, though harder to analyse, view of the improver/improved-by system. By presenting this as a weighted directed network, we were able to use network diagrams and graph theory to summarise our findings in this more complicated setting. We showed how eigenvector centrality gave a similar result to that of our earlier simple ranking of improvement, but was able to give a higher fidelity perspective.

We then went on to show how community detection algorithms could group countries by improver status and which sources they improved. We found a rich cluster and a poorer cluster, with two “bridging” countries which were between the two. We concluded that South Korea would be improved by some countries from the rich cluster and some from the poor, and that Czech Republic was improved by the rich cluster, but acted as an improver to the poor cluster.

But more important than the specific results was the innovative use of vector geometry and graph theory in the study of global trade. These tools have largely been overlooked in economics, and their use here is exploratory rather than definitive. But we feel that the techniques and the visualisations presented here could fruitfully be used in a CGE model as a way of representing model outputs in an engaging and insightful way.

Table 7.4: A list of all sectors in the model. A ✓ indicates the sector is regarded as either services or manufacturing in this chapter.

	Services	Manufacturing
Agriculture		
Air Transport	✓	
Business Services	✓	
Chemicals		✓
Communications	✓	
Construction		
Education	✓	
Electricals		
Financial Services	✓	
Food		
Fuel		
Health	✓	
Hospitality	✓	
Inland Transport	✓	
Leather		
Machinery		✓
Manufacturing		✓
Metals		✓
Minerals		
Mining		
Other Services	✓	
Paper		
Plastics		✓
Private Households	✓	
Public Services	✓	
Real Estate	✓	
Retail Trade	✓	
Textiles		✓
Transport Services	✓	
Utilities	✓	
Vehicle Trade	✓	
Vehicles		✓
Water Transport	✓	
Wholesale Trade	✓	
Wood		✓

7.A Mathematical Appendix

Here we present the detail behind the perturbation process outlined in section 7.3. The purpose of the perturbation process is to rotate the final demand vector (consumption pattern) of the base country, $\mathbf{f}_i$, in the direction of that of a given target country, $\mathbf{f}_j$, by a specified (small) angle θ.

Mathematically, the setting for this process is 35-dimensional Euclidean space, $\mathbb{R}^{35}$. This reflects the structure of the final demand vectors and, ultimately, the fact that the underlying data-set represents economies with 35 sectors. Final demand vectors are expressed in terms of the standard basis $E = \{\mathbf{e}_i\}_{i=1}^{35}$.[12]

The process of vector rotation is a standard mathematical technique consisting of five steps:

1. Use the two final demand vectors to generate a new basis U for $\mathbb{R}^{35}$ via the Gram-Schmidt process.[13]
2. Form the change of basis formula B to map the new basis U into the standard basis E:[14]

$$B = \begin{pmatrix} \uparrow & \uparrow & \dots & \uparrow \\ \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_n \\ \downarrow & \downarrow & \dots & \downarrow \end{pmatrix} \tag{7.7}$$

   The action of the matrix B maps a vector expressed in the new basis U into a vector expressed in the standard basis E. Conversely, the action of the inverse of this matrix, B^{-1}, maps a vector expressed in the standard basis E into a vector expressed in the new basis U.

[12] See mathworld.wolfram.com/StandardBasis.html for a formal description of what a standard basis is. For our purposes it is enough to think of the standard basis as a set of unit-length vectors each pointing in one of the 'directions' of the vector space. Thus, a 3-dimensional standard basis is the set $\{(1,0,0)(0,1.0)(0,0,1)\}$

[13] The Gram-Schmidt process may be thought of as a tool for generating orthonormal bases. Taking an independent set of vectors as an input, the process outputs an orthonormal basis for the space in question (Pursell and Trimble 1991). The input is the linearly independent set

$$\{\mathbf{f}_i, \mathbf{f}_j, \mathbf{e}_3, \mathbf{e}_4, \dots, \mathbf{e}_{35}\}.$$

The first two vectors of the basis U generated by this process, $\mathbf{u}_1$ and $\mathbf{u}_2$, will span the plane spanned by $\mathbf{f}_i$ and $\mathbf{f}_j$.

[14] Any two bases of $\mathbb{R}^n$ can be related by a *change of basis*; this allows us to express one basis in terms of another.

3. Form the Givens rotation matrix $G(\theta)$:[15]

$$G(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 & 0 & \cdots & 0 & 0 \\ \sin(\theta) & \cos(\theta) & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & & \vdots & \vdots \\ 0 & 0 & 0 & 1 & & \vdots & \vdots \\ \vdots & \vdots & & & \ddots & & \\ 0 & 0 & \cdots & \cdots & & 1 & 0 \\ 0 & 0 & \cdots & \cdots & & 0 & 1 \end{pmatrix}$$

   This will rotate vectors expressed in the new basis U by an angle θ in the plane spanned by the first two vectors of the basis U, $\mathbf{u}_1$ and $\mathbf{u}_2$.

4. Form the *final rotation matrix* $R(\theta)$, defined as

$$R(\theta) := B\, G(\theta)\, B^{-1}.$$

   This matrix will map a vector expressed in the standard basis E into a vector expressed in the new basis U, rotate the vector in the plane spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$ by an angle θ, then map the resulting vector back into the standard basis E.

5. Apply the final rotation matrix to the final demand vector of the base country, $\mathbf{f}_i$.

The perturbation process will then produce a new vector

$$\tilde{\mathbf{f}}_i = R(\theta)\mathbf{f}_i,$$

the base country's final demand vector $\mathbf{f}_i$ rotated by an angle θ towards the target country's final demand vector $\mathbf{f}_j$.

Finally, this perturbed vector must be rescaled so that the sum of the components is the same as the original, unperturbed vector:

$$\hat{\mathbf{f}}_i = \frac{\tilde{\mathbf{f}}_i}{|\tilde{\mathbf{f}}_i|}|\mathbf{f}_i|$$

[15] A *Givens rotation* is a rotation in a plane spanned by two coordinate axes. As a rotation is a linear transformation, this can be represented in the form of a matrix acting on the vector we wish to rotate Venkateshan and Swaminathan 2013.

where

$$|\mathbf{v}| := \sum_{i=1}^{35} |v_i|$$

i.e. the sum of the absolute values of the components of the vector $\mathbf{v}$.

Economically, this final rescaling holds total final demand constant throughout the perturbation process: the rescaled perturbed vector $\hat{\mathbf{f}}_i$ and the unperturbed final demand vector $\mathbf{f}_i$ represent the same total demand.

This means that the total demand of the base country's economy is held fixed while the sectoral breakdown varies throughout the perturbation process.

We note that the perturbation process produces monotonic changes to the share of final demand enjoyed by each of the 35 economic sectors, for small angle rotations. This means that our selection of 1 deg for the fixed perturbation distance is not privileged: alternative choices would lead to changes of different magnitudes, but ultimately of the same sign (positive or negative).

Chapter 8

Migration and the global economy*

8.1 Introduction

We described in part I how we would use the model we have now developed to bring additional global economic insight to existing social science models. Our goal was to create the model in such a way that it would be amenable to being "combined" with these other disciplines' models, datasets and findings, in order to shed light on the wider global corollaries of findings which, in their original form, are often either single-country or bilateral. We have now arrived at the first of these expanded analyses.

The work of Dennett (2016) uses a simple proportionality method, discussed in detail in section 8.2.1, to arrive at a time series of estimated global migration flows, using the comparatively well-recorded migrant stock data. Throughout this chapter we will use these flows and denote them as $\bar{\mu}_{ij}$, where the bar denotes a value "from data" and i and j are the source and destination countries respectively.[1]

The original estimates are a time series, but here we only use migration estimates from a single year, thus a time subscript will be implicit rather than stated. Using this time series more explicitly is an interesting avenue for further work.

We begin the chapter with a short literature review on migration flow estimation and the relationship between migration and trade volumes. We then look at the migration estimates themselves, to put some context onto the results of the analyses which make up the remainder. Next we state our two assumptions about the economic nature of migration, ane look at the economic corollaries of each. The first assumption is indis-

---

* This chapter uses the 200-country model throughout. See page 77 for a description of what this means.

1 In this case, "from data" means deriving from work outside this thesis. For practical purposes it means that during the analyses the values will be taken as given. However, this is not the case for the critique of Dennett's methods given in section 8.2.1. Throughout this chapter, we use μ rather than Dennett's original m, in order to distinguish migration flows from imports, m_{ijs}, in the demonstration model.

putable: that migrants bring their consumption dollars with them when they migrate, at least in some form. The form in which this happens will form the topic for the last discussion of the chapter.

The second assumption, perhaps less immediately obvious, is that a larger stock of migrants existing between a pair of countries tends to boost trade between those two countries in a way that goes beyond simple changes in the level of demand. We will examine the evidence for this assumption in section 8.2.2.

8.2 Literature review

This literature review will cover two strands of literature on migration. The first is the literature which precedes Dennett's estimates of migration flows, and seeks to determine both the magnitude of these flows (which, as we shall see, is not well recorded in the data) and the determinants of those magnitudes. The second strand is the literature which relates migration to trade. Here we see that several channels have been proposed to explain the empirical observation that the presence of large migrant stocks from a particular origin country tend to increase trade between that origin country and the host country. We will choose two of these channels to take our own analysis forward.

8.2.1 Migration flow estimates

As Dennett (2016) outlines, data on annual flows of international migrants is surprisingly patchy. Aside from the infrequency of publication of such estimates as do exist, there are a number of large areas in which these statistics are inadequate.

The principal problem is one of a lack of reporting. There is no centralised way of recording when a person moves from country i to country j, leaving aside the question of when someone can be said to have "migrated," as opposed to simply be travelling. Countries in their 10-yearly censuses do a good job of counting countries of origin, but there is no mechanism for counting countries of destination for departing citizens. As Dennett points out, "it is hard to count someone after they have left" (p. 125).

There are some exceptions to this. The OECD has reasonable data on migration inflows. Mayda (2009) for example makes use of this data to estimate the push- and pull-factor causes of international migration flows. She has access to inflow data only, covering 79 origin countries into 14 OECD countries. This data source is not without its problems, although Mayda does not attempt to deal with those that she observes. She notes that the dataset "is supposed to cover immigrant inflows into each of the 14 destination countries from all over the world. However, the sum by country of origin...is not equal to 100% of the total flow into each destination country." (p. 1258).

She estimates a simple linear panel data gravity model, of inflow proportional to population, including both GDP terms and distance, as well as a number of common dummy variables such as a common language and a shared border, dealing with endogeneity issues by simply lagging all dependent variables.

Intra-European immigration and emigration has yet better coverage, and a number of studies have attempted to harmonise these two sets of statistics, and to estimate missing flows. The most well-known of these came out of a University of Southampton project called MIMOSA, which was published by Abel (2010) and in the form of a more widely-read Science Magazine article showcasing the visualisations the team developed to accompany the estimates (Abel and Sander 2014).

Abel's methodology accounts for discrepancies between the reports of origin and destination countries, and attempts to fill missing data using a theory-driven constrained spatial interaction model. The first of these methodologies is more simple.

First, migrants who are reported with unknown origin or destination are distributed according to recorded proportions. Second, the flow data are divided into "reliable" and "unreliable" estimates. The unreliable estimates are discarded, and the reliable ones are harmonised by finding empirically the per-period scale factor which minimises the discrepancy between reliable origin and destination reports, then selecting origin reports only where destination reports do not exist. As he points out, "receiving data are often believed to be of better quality" (p. 804).

The task of filling missing data is done with a spatial interaction model with constrained row and column totals. Additionally to these constraints, the model includes origin/destination ratios of GNI, GDP, total trade value, and the total migrant stock. Finally, dummy variables for mutual Eurozone membership, French-speaking and English-speaking.[2] Although little attention is paid to testing these results in-sample (fits are reported to be "reasonable," but the results are not shown) the methodology is thoughtful and carefully executed.

When the UN published a new dataset of estimated bilateral migrant stocks, Abel and Sander (2014) ran a similar analysis to generate global migration flow estimates. These new estimates were popularised by an attractive circular flow visualisation method which accompanied the publication. An example of this is shown in figure 8.1.

But despite this popularity, Dennett (2016) shows that many European estimates in this global analysis are "some way away from" the latest European flow estimates provided by the IMEM project, an evolution of the original MIMOSA project which Abel himself founded (Raymer et al. 2013). Dennett's estimates are achieved via a simpler methodology than that of Abel, but he nevertheless claims they are "still an improvement on current best estimates" (p. 129).

Dennett's method is far more interested in maintaining basic accounting constraints, than it is with finding relationships between migration flows and macro-level observables. In short, he makes two very simple assumptions in order to move from figures

[2] These mutual dummy variables are equal to 1 when either *both* the sending and receiving countries have the relevant characteristic or when *neither* do.

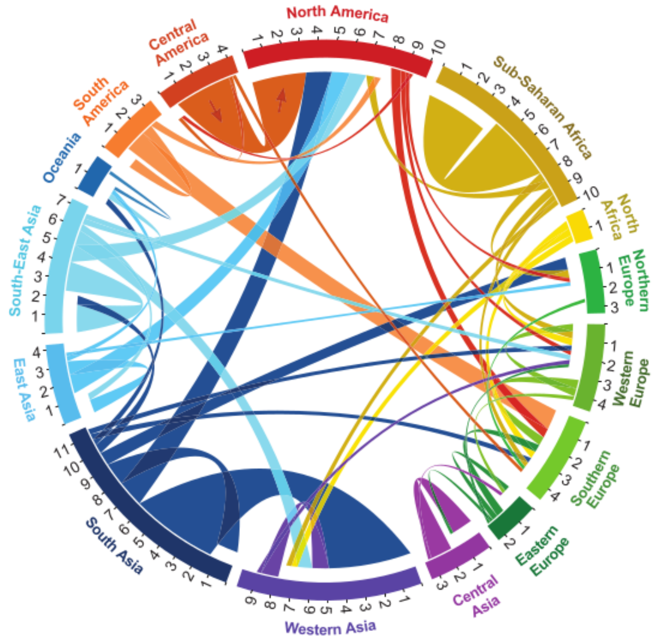


Figure 8.1: A circular flow diagram of Abel and Sander (2014) representing migration flows estimated from UN migration stock estimates.

on migrant *stocks*, which are widely available via census data, to estimates of migrant flow. The first stems from an observation that stocks are highly correlated with flows in Europe as estimated by the IMEM project, with and R^2 of 80 %.[3] This then allows for the equivalence of the two measures when expressed as proportions of global totals.

In order to calculate flow magnitudes, he then requires the estimation of global flow totals. His second assumption is that net migration totals are a good proxy for bilateral migration totals and can therefore be simply aggregated to find a global total number of migrants.[4]

If Dennett's work is extreme in eschewing theory in favour of using data and applying accounting constraints, the other end of the spectrum might be the work of Bertoli and Fernández-Huertas Moraga (2013) (hereafter BF). They describe how the "traditional approach" of regressing only with variables relating to the origin and destination country contrasts with their "multilateral resistance" approach (also described elsewhere in the literature as a network effect) which emphasises the fact that migration decisions depend on the attractiveness/costs of all other destination options facing the migrant.

[3] The possibility that this very strong correlation may be an artefact of the estimation method used by the IMEM project is not considered. It seems to us that this is at least a possibility worth exploring.

[4] Clearly Dennett must confront the fact that one country's immigration is another country's emigration. His aggregation of net migration data can therefore not be simply addition as, by definition, the sum of net migration in the globe must equal zero. No detail is given in his work about how this aggregation is done, but it presumably done by summing positive values and discarding negative ones.

This brings to mind the global modelling approach of this thesis, albeit one based on very strong neoclassical assumptions about the nature of migrants' utility functions.

The approach of BF extends the well-known work on trade of Anderson and Wincoop (2003) to the problem of explaining migration flows. In this latter work, Anderson and Wincoop (hereafter AW) begin with a CES consumer utility function which, in common with all such specifications, introduces a love of variety. This therefore inserts the complete set of all export prices into the demand function for any bilateral trading pair. Using standard neoclassical market-clearing assumptions, AW then derive a familiar-looking gravity equation, with the crucial difference that prices are endogenous, and determined as a function of the complete set of global prices (and, in fact, the complete set of bilateral trade costs.)

This produces a system whereby every ij flow of trade responds to changes in price and bilateral trade costs in every other part of the system, including those not involving i or j. By making greatly simplifying assumptions about the number of elasticities which exist in the global system, they are able to estimate the parameters in their model for the US, Canada and 20 other industrialised countries.

By assuming that there is a particular demand for migration destinations, and that countries "compete" to be the recipient of utility-maximising migrants, BF apply the work of AW using trade as a direct analogy to migration. But on top of this, they add a stochastic term to the migrants' utility to allow for unobserved components of the utility function.[5] These stochastic terms are allowed to be arbitrarily correlated to account for "groups of countries sharing unobservable sources of attractiveness for individuals." This has the benefit of allowing their model an almost unlimited capacity to fit the data, but comes with the usual costs brought by flexibilities of this kind of added complexity and the risk of overfitting the data. They estimate their model using immigration into Spain only, with 52 annual quarters between 1997 and 2009.

Despite the stark simplicity of his approach, Dennett is able to show that his estimates outperform the previous best estimates in comparison both with European IMEM flow estimates, and data on a small number of bilateral flows collated by Kim and Cohen (2010). It is for this reason that Dennett's estimates seem to us to be the best source of global migrant flow estimates although, as with all estimates, results found on the basis of them should be treated as best approximations, not data points. The details of Dennett's estimation set are discussed in more detail in the following section, but first we give a short overview of the literature relating migration and trade.

---

[5] The authors give "cultural proximity" and "civil liberties" as examples of what these unobserved characteristics might be.

8.2.2 Migration and trade

A huge literature is dedicated to exploring the economic factors which determine migration rates between countries. Bertoli and Fernández-Huertas Moraga (2013) give a good overview and describe the recent literature as concerning itself with methodological and theoretical questions. They show how some authors (Beine, Docquier, and Özden 2011; Grogger and Hanson 2011; Ortega and Peri 2013) focus on discrete-choice micro-founded models, and others (Clark, Hatton, and Williamson 2007; Mayda 2009; Pedersen, Pytlikova, and Smith 2008) take a purely econometric approach (i.e. regressions are run to see which observable factors significantly covary with migration rates).

A large number of explanatory variables is employed in these papers, with some being more commonly employed than others. Table 8.1 gives a summary of the variables used in some of the studies the authors include. The most common covariates are perhaps also the most obvious: linguistic distance, the similarity between source and destination countries' languages, and the immigration policy of the destination country. The former of these seeks to explain why people are more likely to move within, rather than between, linguistic blocks such as the English, Russian or Spanish-speaking worlds. The other covariates are either binary variables (Do the countries share a land border? Are they in the Schengen area? Etc.) or single-metric facts about either the source or the destination country (literacy rate, young population share, destination wages, political stability etc.)

For our purposes two things are significant here. The first is that none of these recent papers attempts to describe what effect the *structure*, as opposed to the size, of the destination country's economy has on migration levels. For instance, how does similarity (or otherwise) between the sectoral make-up of the two countries relate to migration? This seems intuitively relevant because workers from a particular sector might find their skills useful in a country which has a high demand for the output of that sector.

This question of economic similarity, beyond just the GDP-measured size of the economies, seems like an odd omission from what is clearly such a plentiful literature, as table 8.1 attempts to demonstrate. It is something which the use of input-output tables, such as we have employed in constructing the demonstration model, makes a very natural question to ask.

The second significant fact about table 8.1 is the presence of migrant stocks as a covariate. Pedersen, Pytlikova, and Smith (2008) look at the idea that having high levels of

Table 8.1: A list of the independent variables used by a selection of recent papers explaining migration flows. Each column represents one of the works cited in Bertoli and Fernández-Huertas Moraga (2013) with an 'x' showing that this variable is used by the author(s).

	Beine	Bertoli	Clark	Grogger	Mayda	Ortega	Pedersen
Linguistic distance	x		x	x	x	x	x
Immigration policy	x	x	x	x	x	x	
Geographic distance	x			x	x	x	x
Colonial relationships	x			x	x	x	x
GDP per capita		x			x	x	x
Average years of schooling			x	x			x
Schengen dummy	x			x		x	
Destination wages	x			x		x	
Relative population size	x						x
Social welfare spending	x						x
Shared land border				x	x		
Young population share			x		x		
Existing stocks	x		x				
Trade totals							x
Cultural similarity							x
Illiteracy rates							x
Political stability							x
Inequality ratio			x				
Source poverty rate			x				
Landlocked source			x				
Refugee/Asylee share				x			
Common currency						x	
Common legislation						x	

trade between two countries tends to increase the stocks of migrants from one country in another. The reverse phenomenon is also very widely studied, as we shall see.

In this chapter, we will seek to extend the idea that trade and migration covary, and see what the corollaries for the source and destination economies this effect has. As the literature reviews in Felbermayr and Toubal (2012) and Genc et al. (2012) demonstrate, the idea that migration between countries is associated with increased trade between them has a long history.

One of the most widely-cited papers in this strand of the literature presents a model seeking to explain why the presence of immigrants in the US have historically been important in increasing bilateral trade flows with the home countries of those immigrants. The author proposes two channels by which this operates: "First, immigrants tend to bring with them a preference for home-country products; and second, immigrants bring with them foreign market information and contacts that can lower the transaction costs of trade." (Gould 1994, p. 303)

He uses the micro-founded gravity model of Bergstrand as a starting point. This uses a neoclassical supply/demand relationship with CES production and consumption functions to formulate trade as a utility-maximisation problem between firms and consumers.[6] To this basic setup, Gould adds endogenous transaction costs associated with trade, which are assumed to "decline with the introduction of foreign market information supplied by immigrants" (p. 304). The equilibrium price of country i's product in country k is inversely proportional to distance-related transport costs, as in a standard gravity model, but also to a non-distance cost, Z_{ik} which is a decreasing function of the migrant stock, M_{ij}. To make the model tractable, it is assumed that $d^2 Z_{ik}/dM_{ik}^2 > 0$.

He chooses a functional form which displays these properties, and estimates its parameters empirically, using total flow data as well as data disaggregated by product category. He finds stronger effects in the consumer goods sector than in the intermediate goods sector, and that exports from the US tend to be affected more than are imports into the US.

Using our input-output approach, we are able to look further into what impact these increased trade flows might have had on both the source and destination countries, as well as any wider effects on the trade network, as per the work of chapter 6 and chapter 10.

[6] See the section of chapter 2 entitled "Trade with microeconomic foundations" on p. 72 for more details on the formulation of the micro-founded gravity model which Gould uses here.

Felbermayr and Toubal (2012) describe two ways in which migration can affect trade. Firstly, similar to the effect explored by Gould, in reducing trade costs by overcoming language, cultural and institutional barriers (we might term this the *familiarity effect*), and second in increasing international demand for goods produced in the source country (which could be termed *consumption similarity*)[7].

[7] The authors of the literature review name these terms the *trade cost channel* and the *preference channel* respectively, but we find these terms less indicative of the underlying economics for our current purposes than our newly-introduced terms.

Table 8.2: The largest and ten smallest percentage emigration flows *per capita* and migration destinations both total and *per capita* in 2010 . Only countries with populations greater than 10 million are shown.

(a) Top 10 largest emigration rates

Country	Emigration (%)
Portugal	0.89
Kazakhstan	0.88
Ukraine	0.57
Greece	0.47
Dominican Republic	0.43
Morocco	0.42
Mexico	0.42
Poland	0.41
Burkina Faso	0.40
Zimbabwe	0.37

(b) Top 10 smallest emigration rates

Country	Emigration (%)
Japan	0.02
Madagascar	0.02
China	0.03
USA	0.03
Brazil	0.03
Nigeria	0.03
Ethiopia	0.03
Saudi Arabia	0.03
Tanzania	0.04
Indonesia	0.04

(c) Top 10 largest migration destinations

Country	Immigrants (000s)
USA	1,548
Russia	505
Germany	404
Saudi Arabia	301
Canada	290
United Kingdom	278
France	274
Pakistan	257
Spain	251
Ukraine	227

(d) Top 10 largest immigration rates

Country	Immigrants (%)
Saudi Arabia	1.11
Australia	1.03
Canada	0.85
Kazakhstan	0.79
Spain	0.54
Côte d'Ivoire	0.53
Belgium	0.52
USA	0.50
Ukraine	0.49
Germany	0.49

8.3 Dennett's migration estimates

We now give a brief statistical overview of the set of estimates, μ_{ij}, which will help us to put the results we find in the remainder of this chapter into context.

The ten biggest and smallest emigration rates[8], as given by the migration estimates, are shown in table 8.2. Since small countries tend, for perhaps obvious reasons, to dominate these statistics, only countries with populations greater than 10 million are shown. This restricts the sample from 214 countries/territories for which the World Bank reports 2010 populations, to the 80 largest countries by population. The smallest

[8] Emigration rate is defined as the number of emigrants per head of the population. As stated above, all figures are for 2010.

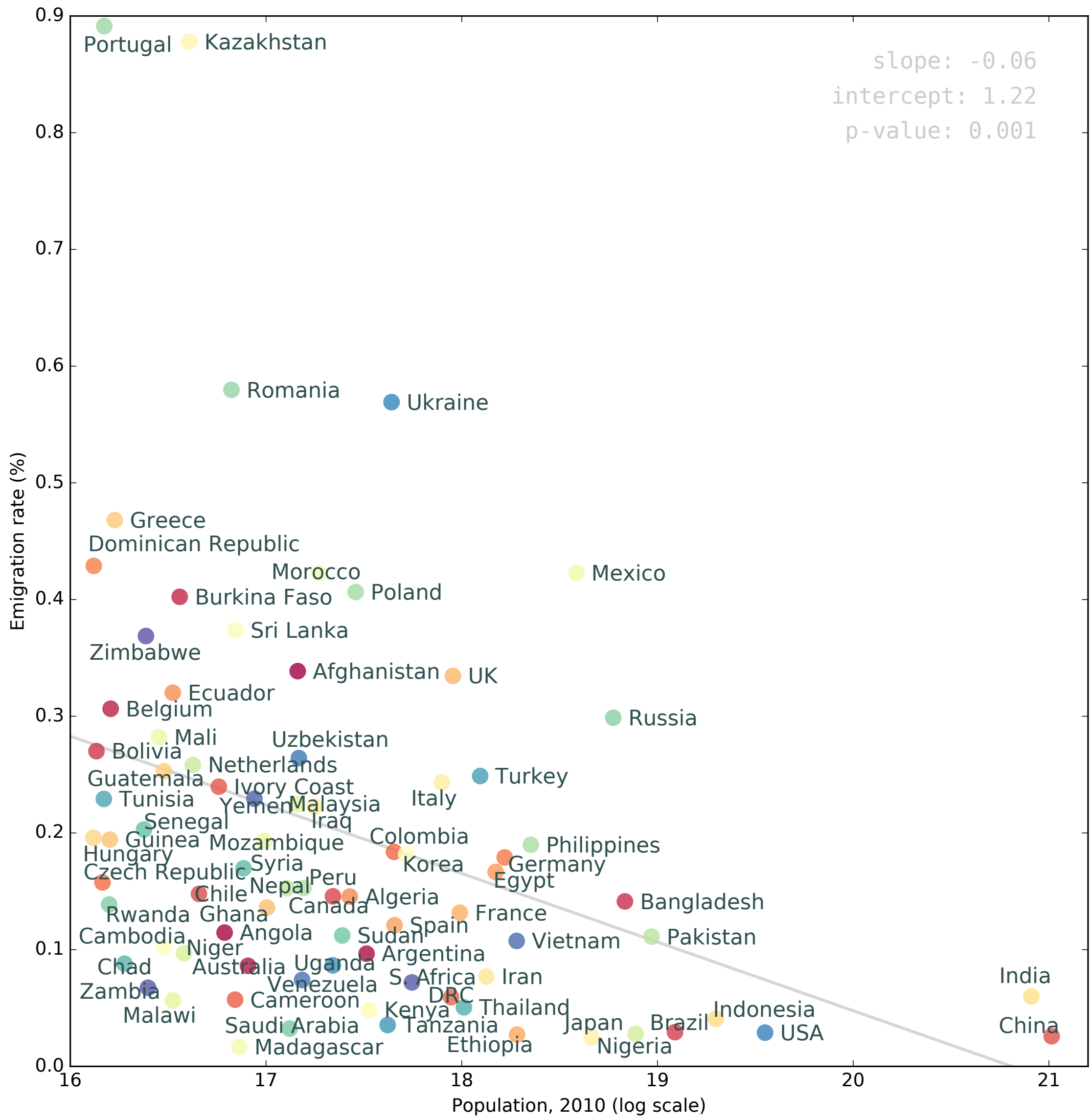


Figure 8.2: The relationship between population size (shown on a log scale) and emigration rates in the 80 modelled countries with a population larger than 10,000,000. The solid line shows the result of a linear fit.

included countries are Tunisia, the Czech Republic and Hungary. The largest excluded countries are Haiti, Guinea and South Sudan.

As listed in table 8.2a, the largest emigration rate among these larger countries is that of Portugal, where nearly 1% of the population emigrated in that year. This could perhaps be put down to the financial crisis which hit that country particularly hard in 2008/09.[9] Kazakhstan showed a similarly large emigration. The four other European countries in the list are the Ukraine, 0.57%, Greece, 0.47% and Poland, 0.40%.

[9] The other European country typically associated with a large post-crisis emigration rate is Ireland, but this has a population of just 4 million and is hence excluded from this analysis. For the record, Ireland had an emigration rate of 0.70% in 2010 and hence would easily have made the top 10 table had it been included.

The smallest emigration rate, shown in table 8.2b, is Japan with just 0.02% of the population emigrating in 2010. Other very small emigration rates were experienced by China, 0.02%, the USA, 0.03% and Brazil, 0.03%. It is perhaps unsurprising that these latter three are amongst the world's largest countries, although India does not show up on this list, suggesting it has a relatively large *per capita* emigration rate for a large country. A more detailed view of the relationship between emigration rates and population size is shown in figure 8.2.

By a huge margin, the largest destination for migrants in 2010, according to these estimates, was the United States with 1.5 million immigrants (shown in table 8.2c), of which almost 500,000 were from Mexico alone. Other large destinations include Russia, with half a million immigrants, largely from Ukraine and Kazakhstan, Germany, 0.4 million, of which over 100,000 were from Turkey, with Saudi Arabia, Canada, the UK, France, Pakistan, Spain and the Ukraine all having between two and three hundred thousand immigrants each.

More surprising perhaps are the *per capita* immigration statistics, shown in table 8.2d. Saudi Arabia is the largest destination for migrants per capita with over 1% of its population being migrants in 2010. There were 300,000 migrants to the country, which has a population of 27 million. They were primarily from India, Pakistan, Egypt and Yemen. It is worth noting that 2010 is the year before a large "Saudization" programme was announced, where Saudi companies were incentivised to increase the number of native workers in their workforce (Maisel and Shoup 2009).

Australia also had over 1% of its population arrive as immigrants that year, primarily from Britain, New Zealand and China. Next on the list of destinations is Canada with 0.85%, primarily from Britain, China and India. Kazakhstan, with 0.79%, appears in both the emigration and immigration lists, with Russia and Ukraine being both the two largest origins and destinations for migration. Spain's 0.54% per capita immigration comes mostly from Morocco, with Britain and the South American countries close behind. Côte d'Ivoire saw the majority of its immigrants coming from Burkina Faso and Mali.

This concludes our overview of the Dennett migration flow estimates themselves. We will now describe how we intend to model the two effects we will use to link these flows of migration to changes in trade between the origin and destination countries and how we can use our global model to give more attention than is usually done on how this affects the economies of those countries, and of the countries in their respective trading networks.

We first take the familiarity effect and show how it can be broken down further into two channels: one treating migrants simply as consumers, thereby affecting the level of final demand, and the other to increased trade volumes. We then proceed in section 8.5 on p. 228 to model the more subtle idea of consumption similarity, whereby migrants are first assumed to consume in exactly the pattern of their source country, and then in a linear combination of source and destination countries.

8.4 The familiarity effect

We can model the familiarity effect, whereby migration flows encourage trade between the source and destination countries, via two separate mechanisms in the demonstration model. The first is a simple magnitude change: the arrival of migrants will presumably boost demand in the destination country and decrease demand in the source country. This will have an attendant increase in import demand for the destination country and an increase in export demand in the source country (as well as in all other countries in the destination country's trade network.) We name this the *demand channel*. For now, we will model this as a scaling of the demand vector with no change in consumption patterns assumed. We lift this restriction in section 8.5.

Secondly, as discussed in section 8.2.2, the source country might become a more preferred exporter for the destination. This we name the *trade channel*. We can model this by adjusting the import propensities of the destination country. We have seen evidence in the literature review of this chapter for a pro-trade effect in both directions. Specifically, Girma and Yu (2002) finds that, in the case of the UK and for migrants outside the Commonwealth, an increase of migrant flows by 10% increases the exports of the destination country to the origin country by 1.6%. Using a more international approach, Felbermayr and Jung (2009) find that a 10% increase in migration between a country pair increases trade by 1.1%. But the direction in which this trade increase happens is not specified since they take the geometric mean of the trade flows in the two directions.

These results and many more have been combined into a meta-analysis by Genc et al. (2012) who find that, though the picture is far from simple, on average, trade in both directions has an elasticity of 0.17 with migration. Thus an increase of 10% in migration increases trade *in both directions* by 1.7%. It is this meta-analysis which we will use here.

For modelling purposes, we must make an arbitrary decision about whether the state of the economy found in the demonstration model in a particular year represents the state before or after the arrival of a particular year's immigrants. For notational simplicity, we will assume that the state of the economy (the level of final demand etc.) is determined *before* the migration flow for that year occurs.

8.4.1 The demand channel

When people leave one country and arrive in another, the very simplest economic effect we can imagine describing is that the products they purchased in their source country will now be purchased in their destination country.

The simplest way to model this "demand channel" is to assume that final demand simply scales with the associated population increase/decrease due to a particular migration flow. This is equivalent to assuming that migrants begin consuming in the same pattern as the households in the destination country.[10]

If we denote with a prime the variables which are changed due to the effect of migration, and denote the population in country i as $\bar{\pi}_i$, with the bar indicating a quantity taken from data, we can calculate the impacts on source and recipient country final demands of the flow of migrants as follows:

$$\begin{aligned} \mathbf{f}'_i &= \left(1 - \frac{\bar{\mu}_{ij}}{\bar{\pi}_i}\right)\bar{\mathbf{f}}_i \\ \mathbf{f}'_j &= \left(1 + \frac{\bar{\mu}_{ij}}{\bar{\pi}_j}\right)\bar{\mathbf{f}}_j \end{aligned} \tag{8.1}$$

where

$$\bar{\mathbf{f}}_i = \begin{bmatrix} \bar{f}_{i1} \\ \vdots \\ \bar{f}_{iS} \end{bmatrix} \tag{8.2}$$

is the final demand vector in country i with elements associated with each sector from 1 to S. The term $\bar{\mu}_{ij}/\bar{\pi}_i$ is the outflow *per capita* of the country of origin. Notice that, because the migrants begin consuming exactly as the households of the destination country, there is no constraint that $\mathbf{f}'_i + \mathbf{f}'_j = \bar{\mathbf{f}}_i + \bar{\mathbf{f}}_j$. In other words, the total level of final demand after the migration will not, in general, be the same as the total level beforehand.

Table 8.3 shows the overall effect of applying equation (8.1) for every migration flow in the dataset for 2010. What this highlights immediately is that gains in GDP (not GDP per capita, note) are broadly negatively correlated with changes to Balance of Trade (BoT). This makes intuitive sense, since when a migrant moves from country i to country j, they stop demanding imported goods from j and start demanding imported goods from i. Generally speaking, countries with generally high import ratios (i.e. a

[10] By "pattern", we mean the ratios in which the output of each of the sectors is demanded.

Table 8.3: The 10 biggest winners and 10 biggest losers from the final demand effect of global migration in GDP terms, measured in millions of $US. Also shown is the change in each country's balance of trade (BoT). Model and migration flows are both using 2010 data.

(a) 10 biggest increases

	BoT	ΔGDP	% GDP
USA	−5,279	55,226	0.42
Canada	−1,409	7,100	0.63
Australia	−1,224	7,016	0.79
Germany	−983	6,507	0.21
France	−703	5,313	0.22
Spain	−950	4,480	0.36
Japan	266	2,224	0.05
UK	−98	1,978	0.09
Italy	135	1,166	0.06
Sweden	−191	1,029	0.22

(b) 10 biggest decreases

	BoT	ΔGDP	% GDP
Mexico	840	−2,270	−0.29
Turkey	265	−787	−0.12
Poland	388	−723	−0.18
Portugal	302	−717	−0.31
Korea	362	−535	−0.07
Romania	204	−486	−0.35
India	226	−233	−0.02
Bulgaria	68	−124	−0.26
Slovakia	101	−106	−0.13
Finland	118	−95	−0.04

reliance on imported goods to satisfy final demand) will do proportionally worse in BoT terms per migrant arriving than those with generally low import ratios.

The biggest winner in GDP terms is, by a huge margin, the USA with a gain in GDP of $55 billion (which represents almost half a percent of GDP). But in fraction of GDP terms, Canada and Australia are bigger winners still (with 0.63% and 0.79% of GDP respectively.) This list should be contrasted with table 8.2c where we listed the top 10 largest migration destinations. Both lists have the USA at the top, but Russia does not benefit as much as its position as the second most common destination for immigrants suggests it might. In line with the reasoning above, this suggests that Russia has a greater reliance on imported goods to supply domestic consumption than does the US. A similar comparison can be made between Canada, which does appear on the top 10 destinations list, and Australia which doesn't. This suggests that the structure of Australia's economy makes it especially able to benefit from the additional consumption of migrants[11].

Mexico is the biggest sufferer from migrant outflow in GDP terms, despite only being the seventh largest in emigration *per capita* terms. It's GDP fell by over $2 trillion, almost a third of a percent of GDP. No other country has an impact on GDP of over a

[11] Or, possibly, from the additional consumption of migrants to its trading partners. Recall that the experiment involves *every* migrant flow in the dataset being resolved in terms of changes to final demand, thus the results include all such *nth*-order effects of migration into and out of all trading partners.

trillion dollars, but Portugal, Romania and Bulgaria are hard hit in percentage-of-GDP terms with 0.31%, 0.35% and 0.26% respectively.

8.4.2 The trade channel

As discussed in the introduction to this section, a flow of migrants may also be accompanied by an increase in trade between the source and destination countries, in both directions.

To model this "trade channel", we can calculate a new set of import propensities by first converting the set of propensities and import demands to a set of trade flows, then adjusting the relevant trade flows (in both directions), and recalculating the import propensities from these newly adjusted flows.[12] We define the trade flows, in terms of the import propensities and the import demands, as:

$$y_{ijs} = p_{ijs} m_{js} \tag{8.3}$$

where p_{ijs} is the propensity of country j to import sector s from country i and m_{js} is country j's import demand for sector s. In order to apply the elasticity of 0.17 of Genc et al. (2012) we must first determine the change in migration flows between the year in question and the previous year, which we will denote $\Delta\bar{\mu}_{ij}$. Then from the definition of an elasticity:

$$\frac{\bar{\mu}_{ij}}{y_{ijs}} \frac{\Delta y_{ijs}}{\Delta\bar{\mu}_{ij}} = 0.17 \tag{8.4}$$

$$\implies \Delta y_{ijs} = y'_{ijs} - y_{ijs} = 0.17 \frac{\Delta\bar{\mu}_{ij}}{\bar{\mu}_{ij}} y_{ijs} \tag{8.5}$$

$$\implies y'_{ijs} = \left(1 + 0.17 \frac{\Delta\bar{\mu}_{ij}}{\bar{\mu}_{ij}}\right) y_{ijs} \tag{8.6}$$

When $\bar{\mu}_{ij} = 0$, i.e. there is zero migration from i to j, a relatively common occurrence among the estimates, the term in equation (8.6) is undefined and we simply set $y'_{ijs} = y_{ijs}$.

Notice that, for now at least, we assume the impact on trade flows is identical across all sectors. Improving on this assumption in a non-arbitrary way might be an interesting avenue for future research.

We can then use equation (8.6) along with equation (8.3) to calculate a new set of import propensities using the definition of an import propensity from section 3.3 as a

[12] See chapter 3 for a complete description of all the terms used in the mathematics of this section.

share of import coming from a particular country:

$$p_{ijs} = \frac{y'_{ijs}}{\sum_k y'_{kjs}} \tag{8.7}$$

where all y' have been adjusted for the trade channel of the familiarity effect. Notice that since a particular set of $p_{\bullet js}$ is constrained to sum to unity, a change in such final demand will result in the recalculation of all others.

Thus, each migration flow affects the import propensities and, via equation (8.3), the trade flows between some subset of the modelled countries.[13] To this new set of import propensities we can now add the adjusted final demand in the source and destination countries, calculated via equation (8.1). We can then use both these changes to assess the impact of each migration flow on the GDP and GDP per capita of each country in the model, as well as on total production in the model as a whole.

Table 8.4 shows the biggest increases and decreases in emigration between 2009 and 2010 as a fraction of the country's 2010 population. Results are shown only for the 40 WIOD countries, although emigration figures were calculated across all countries in the estimate data set. These results will be important when we analyse the effect of the trade channel below. The biggest reduction in emigration in 2010 is in Ireland, where there were fewer migrants than in 2009 equal to 0.05% of the population. Estonia and Cyprus are next on the list with 0.030% and 0.027% respectively. Luxembourg shows the biggest increase as a proportion of its population over the same period, followed by Portugal and Romania.

The effects of applying the trade channel to each country in the model are shown in table 8.5. The largest winner from this effect, shown in table 8.5a, is China by a wide margin, which gains $2 billion, or 0.06% of its 2010 GDP. But far more significant from a percentage perspective, are Belgium and the Netherlands, both with over a billion dollars of GDP increase from the trade channel, or 0.22% and 0.14% of GDP respectively. But none of these gains is huge in percentage of GDP terms.

The really significant results are among the GDP losses, shown in table 8.5b. Two countries, the Czech Republic and Ireland, both suffer by more than one percent of GDP from the trade channel in 2010, with absolute reductions in GDP of $3 billion and $2.4 billion respectively. Both these countries are high on the list of emigrant

[13] Precisely *which* subset, depends on the nature of the wider trade network the source and destination countries are part of. In many cases in fact, a change in one import propensity will affect the trade flows between *every* pair of countries in the world, although some of these changes will be very small.

Table 8.4: The ten biggest increases and reductions in emigration between 2009 and 2010 across all destinations. Changes are shown as a percentage of the country's 2010 population. Results are only shown for the 40 countries of WIOD, but emigrations were calculated across all countries in the data set.

(a) The ten biggest increases

Country	Δ Emigration (%)
Luxembourg	0.037
Portugal	0.034
Romania	0.031
Malta	0.030
Bulgaria	0.020
Belgium	0.018
Turkey	0.006
Slovenia	0.006
Austria	0.005
Mexico	0.005

(b) The ten biggest reductions

Country	Δ Emigration (%)
Ireland	-0.051
Estonia	-0.030
Cyprus	-0.027
Czech Republic	-0.026
Poland	-0.025
Lithuania	-0.022
Latvia	-0.021
Finland	-0.008
Slovakia	-0.007
Russia	-0.005

reductions in table 8.4b. Russia, Japan, Germany and Poland all have GDP losses of over a billion dollars, but none of these is as significant in percentage terms.

8.4.3 A combined familiarity effect

We now combine the demand channel and the trade channel into a single familiarity effect. This takes account of the magnitude changes of the demand channel, and the increased bilateral flows of the trade channel. To combine them, we simply apply both effects for each migration flow then recalculate the model once all the coefficient adjustments have been made. The results of this are shown in table 8.6.

The three biggest GDP gains, shown in table 8.6a, are experienced by the USA, Canada and Australia, the same three as at the top of the table for the demand effect in table 8.3a. The gains seen with the combined effect are very slightly increased for all three. It can be seen from the fact that none of these three countries is in the top 10 GDP gains from the trade channel, in table 8.5a, that the combined effect on GDP of the two channels is not equal to the sum of the effects. This is to be expected, since the countries are interlinked in a complex manner via the trade model, and therefore changes in one economy affect all others in a way that is not immediately straightforward to predict[14].

[14] Although because of various monotonicities in the model's equations, changes in demand in one country can only ever lead to changes in the same direction in others.

Table 8.5: The biggest gains and losses from applying the trade channel of the familiarity effect, governed by equation (8.6). Only results for the 40 countries of WIOD are shown, and the trade channel was only calculated across these countries. GDP changes are shown in $US millions.

(a) Ten biggest GDP gains

	ΔGDP	%
China	2,245	0.06
Belgium	1,124	0.22
Netherlands	1,075	0.14
South Korea	792	0.11
Spain	647	0.05
UK	469	0.02
Italy	419	0.02
Mexico	353	0.04
Austria	171	0.04
Sweden	94	0.02

(b) 10 biggest GDP losses

	ΔGDP	%
Czech Republic	−3,057	−1.66
Ireland	−2,420	−1.00
Russia	−1,864	−0.16
Japan	−1,655	−0.04
Germany	−1,496	−0.05
Poland	−1,125	−0.28
Indonesia	−765	−0.15
Finland	−707	−0.33
Brazil	−498	−0.03
Slovakia	−466	−0.58

Spain, France and Germany then follow, as they do in the demand channel, but with Spain leapfrogging France and Germany into fourth place. This reflects the fact that Spain appears in biggest trade channel gains, where France and Germany do not.

Japan, Italy and Sweden fail to appear in the combined table. Japan has one of the biggest losses from the trade channel, but both Italy and Sweden are among those benefiting from the trade channel. These latter two only fail to make the top 10 because Belgium, the Netherlands and China have such strong gains from the trade channel.

The two biggest sufferers from the combined familiarity effect, shown in table 8.6b, are the same as those from the trade channel, shown in table 8.5b. They are the Czech Republic and Ireland. Mexico is actually among the 10 biggest increases due to the trade channel, but is such a big sufferer from the demand channel that the costs outweigh the benefits. Poland suffers from both effects, as does Finland, explaining their positions among the worst combined sufferers.

Turkey has the second-worst decrease in GDP due to the demand channel, so is therefore among the worst affected by the combined effect, despite it having among the largest increases in emigration between 2009 and 2010, as shown in table 8.2a. Together with Mexico and Portugal, which also had large increases in emigration in that year, Turkey shows a large improvement in its balance of trade, set against its large reduction in GDP. This shows that the various effects of migration tend to work in opposite directions, with gains (reductions) in GDP being accompanied by commensurate reductions (gains) in balance of trade. Notable exceptions to this are China, the Nether-

Table 8.6: The biggest gains and losses from applying both channels (demand and trade) of the familiarity effect. Only results for the 40 countries of WIOD are shown. GDP changes are shown in $US millions.

(a) Ten biggest GDP gains

	BoT	ΔGDP	%
USA	−5,588	54,917	0.42
Canada	−1,667	6,842	0.60
Australia	−1,669	6,571	0.74
Spain	−301	5,129	0.42
France	−888	5,128	0.22
Germany	−2,483	5,007	0.16
China	3,536	2,821	0.08
UK	372	2,448	0.12
Netherlands	1,117	1,890	0.25
Belgium	1,083	1,849	0.36

(b) 10 biggest GDP losses

	BoT	ΔGDP	%
Czech Rep.	−3,019	−2,991	−1.63
Ireland	−2,397	−2,247	−0.92
Mexico	1,194	−1,916	−0.24
Poland	−741	−1,851	−0.46
Russia	−1,795	−1,248	−0.11
Turkey	176	−876	−0.14
Finland	−591	−804	−0.38
Indonesia	−595	−769	−0.15
Portugal	390	−628	−0.27
India	−120	−579	−0.05

lands and Belgium which all see large gains in balance of trade accompanying their increases in GDP. On the other side are the Czech Republic, Ireland and Russia which all have large reductions in balance of trade as well as having reductions in GDP.

This approach to modelling the effect of changing populations allows us some novel ways of thinking about how migration interacts with the wider economic system. We have seen how migration has conflicting effects on GDP via the demand channel, whereby emigrants take their expenditure with them to their destination country, and via the trade channel, whereby a source country's trade is boosted bilaterally with a destination country. We have also seen how these opposing forces have varying effects on balance of trade and GDP, depending on a country's position in the global network, and its relationships in both a trade and a migration sense with the other countries in the model.

8.5 Consumption similarity

In section 8.4 we treated migrants' consumption patterns as being identical to those of the population of the destination country. This allowed us to simply scale the final demand vector according to migration as a proportion of the population.

A more interesting way of modelling this process might be assume either that migrants continue to consume in the same pattern as the population of the source country, or as some linear combination of the two patterns. By making these more subtle adjustments to the final demand vector as a result of migration, we can compare the effect on the economy of the destination countries and, by some measure, assess the extent to which a particular country's demand vector is changed by the immigrants it hosts. In this way, we are making a link to the vector rotation work of chapter 7.

Early work on this subject by Wallendorf and Reilly (1983) inspected the contents of the rubbish bins of Mexican Americans in an attempt to test the theory that migrants consume in a pattern mixed between the source and the destination countries. They in fact found that migrants consumed according to their own unique patterns. In the context of Greenlandic immigrants to Denmark, the work of Askegaard, Arnould, and Kjeldgaard (2005) finds evidence both that consumption patterns matched the destination country and that new consumption patterns were developed. Karamba, Quiñones, and Winters (2011) find, in a food context, that migrants in Ghana did not change their consumption expenditure or consumption pattern from that of the source country. The picture in the literature is therefore a mixed one, between no consumption pattern adjustment and perfect consumption pattern assimilation. It therefore seems reasonable to take an approach some way between these two extremes.

We will therefore assess the impact on the destination country of migrants variously adjusting their consumption patterns. The patterns will vary between completely retaining the source country patterns, to complete consumption-pattern assimilation. In order to simplify the analysis we will focus on the three largest recipients of immigrants. Since adjusting consumption patterns requires knowing the patterns for the source country, only migrants from the 40 WIOD countries will be considered. With this restriction, the largest three are the USA, Germany and Canada, with 0.9 million, 0.3 million and 0.2 million immigrants respectively in 2010.

To make the adjustment to the final demand vector, we will use the simplest possible extension to equation (8.1) where the adjustment term is a linear combination of the source consumption vector and that of the destination. Thus, the adjusted consumption

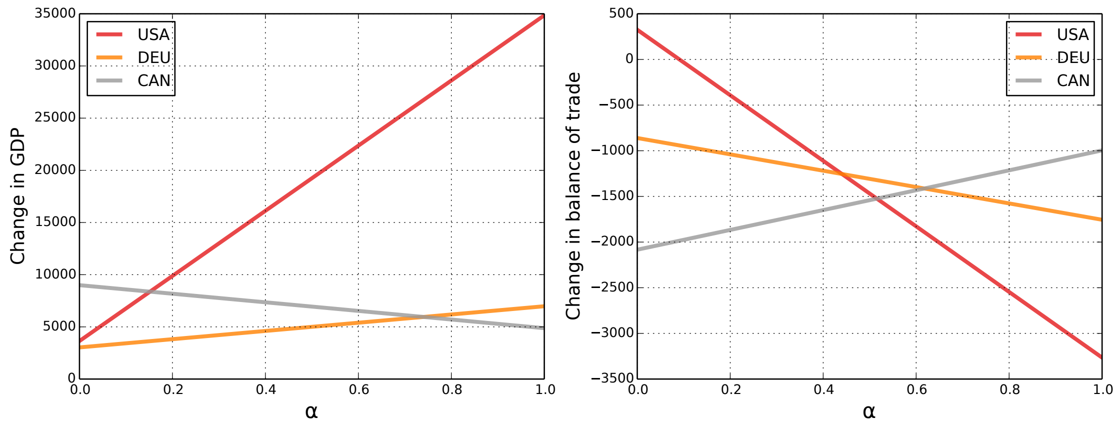


Figure 8.3: Changes in GDP and balance of trade for the three largest recipient of WIOD-country migrants in 2010 for various levels of the linear combination parameter α in equation (8.8).

vector at destination j due to a migration flow from source i is given by:

$$\begin{aligned} \mathbf{f}'_i &= \left(1 - \frac{\bar{\mu}_{ij}}{\bar{\pi}_i}\right)\bar{\mathbf{f}}_i \\ \mathbf{f}'_j &= \bar{\mathbf{f}}_j + \frac{\bar{\mu}_{ij}}{\bar{\pi}_j}\left[\alpha\bar{\mathbf{f}}_j + (1-\alpha)\bar{\mathbf{f}}_i\right] \end{aligned} \tag{8.8}$$

where, as before, a bar indicates a value taken from data, $\bar{\mathbf{f}}$ is a vector of per-sector final demand values from WIOD. The linear combination parameter α controls the extent of consumption-pattern assimilation. With $\alpha = 0$ we have no assimilation at all, and $\mathbf{f}'_j = \bar{\mathbf{f}}_j + \frac{\bar{\mu}_{ij}}{\bar{\pi}_j}\bar{\mathbf{f}}_i$. With $\alpha = 1$ we have perfect assimilation and equation (8.8) reduces to equation (8.1) and we are back to our earlier simple final demand scaling. Figure 8.3 shows the results of taking this parameter across the whole range from no assimilation to perfect assimilation.

The three countries chosen have markedly different experiences both in their sensitivity to α and in their GDP vs. balance of trade changes. At $\alpha = 1$ Canada and Germany have very similar improvements to GDP, both around \$5 billion, and the US has a hugely greater improvement at around \$35 billion. These figure are broadly in agreement with the demand channel discussion of section 8.4.1, given in table 8.3[15].

But when $\alpha = 0$, and migrants continue to consume in the exact pattern as they did in the source country, the picture looks extremely different. Here, Canada has a small advantage, increasing its GDP gains to just under \$10 billion, and Germany stays broadly the same, falling to just under \$5 billion. But the US falls dramatically by more than \$30 billion to just under \$5 billion, almost identical to the gains of Germany. This

[15] The numbers do not match precisely since that analysis used migrants from every country in the data set, where this uses only migrants from the 40 countries of WIOD.

huge difference suggests that a large proportion of the huge gains shown by the US in the discussion of the familiarity effect of section 8.4 is contingent on the migrants who arrive adapting their behaviour to match that of their new compatriots. Without this effect, the US is in no better a position than Canada or Germany.

Between the extrema of α is what appears visually to be a straight line continuum of GDP benefits. In fact this relationship is not quite linear. For the US, the relationship is increasing in α, with $\frac{\partial^2 \text{GDP}}{\partial \alpha^2} = 0.6$, and for Canada it is decreasing, with $\frac{\partial^2 \text{GDP}}{\partial \alpha^2} = -0.6$. For Germany the relationship is very close to linear.

The picture for balance of trade shows even more variation than that for GDP. At $\alpha = 1$, all three countries show a reduction in the balance of trade, which is a natural result of losing consumers in the foreign market and gaining domestic consumers. The US suffers most severely with over \$3 billion of losses, with Germany just over \$1.5 billion and Canada at \$1 billion. But at the other end of the scale, with $\alpha = 0$, the picture is completely reversed. Canada is now the worst off, with a small decrease in its situation to around \$2 billion of losses. Germany improves slightly to around \$1 billion. But for the US the difference is again extremely marked, with the balance of trade now changing positively by around \$300 million. At around $\alpha = 0.5$, the three countries are affected similarly in balance of trade terms. This suggests that having migrants who behave like Americans is particularly bad for balance of trade in the US, but having migrants who continue to behave like the population of their source countries may actually be beneficial to the US balance of trade. Conversely it is better for Canada's balance of trade if immigrants behave like Canadians.

8.6 Conclusion

In this chapter we have used the migration flow estimates of Dennett (2016) to analyse some proposed mechanisms via which migration may be linked to the wider economic context.

We began by hypothesising that migration affects trade via a familiarity effect, whereby not only did the migrants take their consumption dollars from the source to the destination country, but their very presence in the destination increased bilateral trade in both directions. We "applied" each of the migration flows in the data set and saw that the results were very mixed across the modelled countries. Some countries, notably the USA, Canada and Australia, experienced large increases in GDP, and others, the Czech Republic, Ireland and Mexico suffered large decreases. The correlation between GDP changes and balance of trade changes was found not to be straightforward. We also found that the two channels by which the familiarity effect was proposed to work, the demand channel and the trade channel, had different levels of importance in different countries.

We then relaxed the assumption that migrants consumed in the same patterns as people in the destination country and studied the effect of changing migrant consumption patterns in three large migrant destinations. Here we found that the results from the familiarity effect were extremely contingent on how the migrants' consumption patterns changed from source to destination country. We also found that changes in GDP were broadly negatively related to changes in balance of trade.

Once again, as with much of the analyses in this thesis, this very brief analysis is intended to show the flexibility of our model, and the kind of wider economic context which can be hinted at by using a global model of trade, rather than to be a complete or in-depth study of the relationship between migration and the global economy.

The global nature of our model allows us to see the gains from migration as well as the losses the must occur elsewhere in the network if trade is to increase between a particular pair of countries without a commensurate increase in final demand. In other words, we are able to see how causes which result in the favouring of a particular country as a trade partner must inevitably have an effect on those countries which are *disfavoured* by those same causes. This is the kind of wider economic context of which we are hoping to give a flavour throughout this thesis.

Chapter 9

International Security and the global economy*

9.1 Introduction

The second of our expanded analyses demonstrates how the model we have developed can bring a new source of global economic inference to the much-studied area of international security.[1]

The link we will draw between the fields of security and economics comes via the changing of military expenditure within one state in response to perceived threat levels from the other states in the model (Collier and Hoeffler 2002; Nordhaus, Oneal, and Russett 2012; Böhmelt and Bove 2014). Specifically, we use a recent measure of the threat between two countries introduced by Baudains, Fry, et al. (2016). In work by the same authors, this threat measure is found to be a predictor of military spending and it is this relationship that is key to the work of this chapter.

We develop a very simple mechanism by which this increased expenditure affects final demand and trade relationships and apply these changes to our model. This then allows us to draw some global inferences from a set of country/country relationships. As we show in section 9.3, the measure of threat we use from Baudains has some unusual and unexpected properties. This inevitably colours the results of this chapter and means, as with all the analyses of this thesis, that we are presenting a series of proposed *methods* for linking these fields, and giving global context to security questions. Performing these analyses with a more convincing threat measure, and possibly with an expanded

---

* This chapter uses the 200-country model throughout. See page 77 for a description of what this means.

[1] See part I and the introduction to chapter 8 for a description of what we mean by the term "expanded analysis", and what our goals are in performing such analyses.

version of the demonstration model which more closely approaches a fully-specified CGE model, is left for future work.

We first give a brief overview of the security literature generally, and how it relates to economics more broadly. We find global studies far more rarely than country-specific studies, but some prominence is given to those global studies which *do* exist despite their comparative rarity.

9.2 Literature Review

Despite being very widely studied, the question of how the government spending on defence affects the national economy remains a deeply contentious issue. Studies since at least the 1980s have attempted to prove the existence of both a positive knock-on effect of military expenditure and, just as often, of a "peace dividend" whereby it is *reductions* in military expenditure which have a positive effect on the economy.

With many regression studies employing essentially the same data, the focus has rather been on how to model the channels via which defence spending affects the wider economy. For our purposes, it is relevant to survey these proposed mechanisms and, wherever possible, to find studies which have a global, rather than a country-level, focus. This latter challenge will prove difficult.

The literature review is divided up as follows. We first give an overview of the very voluminous regression literature on military spending and economic growth. We then go on to look at the same question studied using computable general equilibrium (CGE) models. We give a brief overview of some more global approaches, before turning our attention to the literature on measuring military threat. It is this last strand of literature that will lead us to Baudains' estimates of dyadic threat which we will examine in more detail in the following section.

9.2.1 Military spending and growth

The literature attempting to find a relationship, positive or negative, between military expenditure (hereafter milexp) and economic growth is large enough to have been exhaustively surveyed several times since the 1980s. The most well-renowned of these, and perhaps the most opinionated, is Dunne, Smith, and Willenbockel (2005). We will refer to their opinions of much of this work throughout this section, referring to them from here on as DSW.

A more rigorous and more recent survey, with some of the hallmarks of a proper meta-analysis, is undertaken by Alptekin and Levine (2012), who divide the literature usefully into three strands.

The first of these strands is due initially to Feder (1983) in a context which is not related to milexp. Instead, Feder's model is of an economy with two sectors, exports and "non-exports," which are assumed to have different marginal factor productivies.[2] This difference allows Feder to explain economic growth not only through growth in

[2] Marginal factor productivity is defined as the number of units of each factor of production, labour and capital, and sometimes intermediate goods, required to produce one additional unit of output.

capital and labour supply, but also through re-allocation of resources from the less-productive to the more-productive sector. His key modelling assumption, which will go on to be relevant to the milexp literature, is that these marginal productivities are in the same ratio between sectors for each factor. More formally, if one sector's production function is denoted by F and the other by G, and partial differentiation with respect to either capital or labour is denoted with a subscript then:

$$(G_k/F_k) = (G_l/F_l) = 1 + \delta \qquad \text{(Feder eq. 3, p. 61)}$$

The problem of determining which sector is more productive is then reduced to the estimation of the single parameter δ and assessing whether it is greater than or less than zero.

This formulation is used by Ram (1986), but with the two sectors being government and "nongovernment." He extends Feder's formulation by allowing the output of the nongovernment sector to depend explicitly on the output of the government sector, adding an elasticity parameter, which Ram denotes θ. It is precisely this extended version of Feder's two-sector model which is then first applied to milexp explicitly by Biswas and Ram (1986). Here the two sectors are now military and civilian, and the estimation of the two parameters amounts to assessing which sector is better at using factors of production (δ) and how milexp directly affects civilian output through spillovers (θ). This approach was then very widely adopted, and Ram himself provides a survey of its use (Ram 1995).

Despite the appealing simplicity of this approach, and its ease of estimation from data, DSW are unequivocal in their rejection of it. They begin their critique by noticing that the assumption that for the outputs of the two sectors to be comparable they must be measured in monetary values not quantities. This they describe as an "implicit price normalisation" but one in which prices are not explicitly dealt with. Comparisons of the marginal factor productivities between sectors "necessarily depend on the prices used in the evaluation of sectoral outputs" (p. 454). They use this observation to go on to point out the theoretical flaws in the model, directly citing from many scholars who have used it without question.

The second of Alptekin and Levine's three literature strands uses a simultaneous-equation framework, and begins with Deger and Smith (1983). Their model allows for a direct effect of milexp on growth "through resource mobilization and modernization effect" (p. 340) and an indirect effect assuming that milexp uses money which could otherwise be invested/saved. It is also the first to make milexp itself endogenous.

Some seemingly very serious theoretical issues in making milexp endogenous are waved away without much analysis. For example, countries are assumed to have identical propensities towards milexp. The authors accept that "the strategic environment faced by the countries in the sample varies very widely" but are happy to account for this with the inclusion of two dummy variables, "one for oil-producing states and one for states involved in external wars." Milexp is also assumed to be proportional to income per capita "since a large military establishment has relative status connotations" (p. 343).

These difficulties aside, the Deger–Smith model has three equations, for growth, savings and military expenditure. These are jointly estimated using a technique which accounts for the simultaneity and it is found that milexp has a small modernisation effect on growth, which is outweighed by a large negative effect on the saving rate. Despite the appealing plausibility of the outcome, DSW criticise the specification for its simplistic treatment of endogenous military spending, pointing out that the error term includes "threats against which military spending is effective and possible measures of the potential for rent-seeking behaviour by the military industrial complex" (p. 452). The absence of threat-level comes in for particular scrutiny, since it may also enter the output equation, thus rendering any empirical findings moot.

The third strand demarked by Alptekin and Levine (2012), and that favoured by DSW seeks to align the security/growth literature with the economic growth literature more broadly. This is done via the use of a "Barro-type growth model" or an "augmented Solow growth model." We will give these growth models a very brief overview here.

The Barro growth model (Barro 1990) is characterised, and made distinct from the standard Solow model, by constant returns to capital and an endogenously determined savings rate. The Solow model assumes diminishing returns to capital which, for Barro and Sala-i-Martin (1997), led to a problematic result that "the economy will eventually converge to a steady state with zero per capita growth" (p. 61)[3] The Barro model replaces this with constant returns to capital, justified qualitatively by allowing capital to include "human capital." This was important to Barro because, in a diminishing-returns world, poor countries should be catching up in income terms to rich countries, something that, at the time at least, was not happening. The addition of a representative consumer who optimised between saving and consumption over time allows the Barro model to respond to changes in the interest rate, tax, and other variables affecting consumer behaviour. In contrast the Solow model has no consumer behaviour at all.

[3] The ideas that either we may then have *been* at the point where diminishing returns to capital have begun to bite, or that we now *do* seem to be in a world of systemically low growth seems not to have troubled Barro and Sala-i-Martin. For them, limited growth was merely a theoretical problem to be overcome.

An alternative growth model which is also popular in the security literature is the "augmented Solow model" of Mankiw, Romer, and Weil (1992). These authors reject the constant-returns hypothesis, finding empirical evidence that, once other variables are accounted for, countries do indeed tend to converge toward one another over time. They add human capital to the Cobb-Douglas production function of Solow, and allow the savings rate to differ between savings invested in human capital, and those invested in physical capital. This allows them to dismiss many of Barro's original complaints about the Solow model.

A well-known use of Barro-type models is given by Aizenmann and Glick (2006). Their production function has a Cobb-Douglas mix of capital/labour ratio and government spending on infrastructure (non-military). The also include an output cost of "threat posed by foreign rivals' actual or potential hostile action" (p. 137). Milexp is modelled for by assuming it acts to reduce this threat cost. This allows them to distinguish between corrupt military spending, in the absence of foreign threats, which is found to reduce growth, and genuine military spending, when foreign threats are present, which increases it.

A similarly well-cited study which uses the augmented Solow growth model is that of Yakovlev (2007). In a paper which compares the augmented-Solow and Barro models, Yakovlev finds that the augmented Solow performs better when taken to the data using standard panel data estimation techniques, and that military spending general reduces the rate of growth. This effect is found to be dampened when the country is a net arms exporter, although no qualitative argument is proposed for how this mechanism might operate.

9.2.2 General equilibrium approaches

The most widely-used approach when studying the economy in a multi-sector context, beyond the two-sector frameworks outlined above, is the general equilibrium modelling described in detail in the literature review of this thesis, in the section beginning on p. 61. In this context, these are usually single-country models, with a fairly simplistic view of the rest of the world.

Athanassiou, Kollias, and Zografakis (2002) use an existing computable general equilibrium (CGE) model of Greece to study the effect of military spending on the wider economy. The details of the model are provided by a pre-existing social accounting matrix from the late 1980s.[4] Once the model is built, the approach is a simple one. The

[4] This highlights one of the often-cited difficulties with CGE models, and with all data-driven models of this sort, including the input-output models we use in this thesis. The data requirements are extensive

authors obtain government military spending from another data source, and imagine what would happen to the economy if that money were spent elsewhere.

The three non-defence sectors they use for this purpose are education, health and “other government” which, although not further specified, presumably includes capital expenditure, public sector wages, transfers and debt-servicing costs among other things. They add the military expenditure quantities first to each sector proportionally, then to each sector in turn, measuring the changes in the various macroeconomic variables which their well-specified CGE model allows. They conclude, among many other points, that milexp favours unskilled, “non-urban” employees and is directed towards “the more traditional sectors” by which they mean those sectors towards the primary input end of the economy.

A similar approach is taken by Yang et al. (2015) for the case of South Korea. The innovation here is that they use an idea due to Dunne, Smith, and Willenbockel (2005), described above, of separating the effects of military spending into three “channels.” These are demand, supply and security itself.

The demand channel is a Keynsian-style demand multiplier and an increase in employment if there is spare capacity in the labour market. The supply channel is represented as a possible opportunity cost to other sectors, as investment in the military crowds out investment and expenditure elsewhere. Additionally, positive spill-overs in the form of knowledge exchange and human capital development are allowed for, meaning that the supply-side effect can have either a positive or negative sign.

Finally, the security channel looks at milexp *per se*, rather than as a generic economic expenditure as do the other two channels. There are a number of proposed phenomena at work here. The additional feeling of security engendered by military expenditure is assumed to lead to a more optimistic business environment: “military expenditures can enhance incentives to accumulate capital and produce more output” (Yang et al. 2015, p. 598). This is the idea endorsed by Thompson (1974) who describes defence as being neither a consumer nor a producer good, but rather “a good produced to protect the ownership of other goods from foreign takeover” (p. 755). Another mechanism, due to Aizenmann and Glick (2006), is that milexp takes two forms. Expenditure due to genuine external threats, which has a positive effect on the wider economy, and that due to “rent-seeking” or corruption, which is an inefficient use of resources and can result in wasteful arms races.

---

and are often published on an infrequent basis. Although this Greek study is an extreme example of this, the general phenomenon is very widespread.

Using an input-output table from 2009 with 28 industries, three factors of production and separate civilian and military government sectors, Yang et al. analyse the effect of an increase in military budget matching a 2009 defence plan of the Korean government, measuring changes from a baseline scenario up to 2040. This is done first by assuming the money is taken from other parts of the budget, and then by assuming an increase in taxes of various kinds in a number of scenarios. They broadly find that milexp funded by cuts elsewhere in the economy is very slightly detrimental to GDP growth, but that funded by tax increases is slightly beneficial. With their CGE model they are further able to forecast the effect of these various scenarios on private consumption, social welfare and labour demand.

Despite the overall impressive nature of these CGE approaches, they remain essentially national-level analyses with little attention given to the multiplier effects available via increased trade, improved economic output in trading partners etc. We were surprised to find little mention of military expenditure issues in the extensive GTAP modelling literature, an obvious candidate for studies of a global nature.[5] A very small number of papers have any kind of global approach to military spending, and we now turn to those few which do.

9.2.3 Other global approaches

Beyond the cross-section regression approaches of, for example, Barro and Sala-i-Martin (1997) and the literature which follows that study, there are few attempts at taking a global perspective on military expenditure. Heo (1998) uses an interesting non-linear specification, with its roots in a theoretical model of production, by the resulting specification still models only within-country effects.

An exception is the geographical work of Shin and Ward (1999), who use spatial dependence measures to look at regionally specific explanations for the defence-growth trade-off. They describe vividly the problems with a per-country approach to military expenditure:

> it seems trivial to argue that a country's military spending and economic productivity are, in part, a function of its interactions with other states. Yet, few econometric models in the defense-growth trade-off literature attempt to evaluate or control for these effects (i.e., geography). By default, this omission leaves one to assume incorrectly that either (1) countries are spatially independent from one another or (2) interactions between states are homogeneous and isotropic. (Shin and Ward 1999, p. 794)

[5] For a complete discussion of the GTAP model and its surrounding literature see the section of the literature review starting on p. 64.

By using standard spatial-regression techniques which look for contiguous areas with similar characteristics, they find that economic growth is positively correlated with both own-country growth and neighbouring-country growth and that there is a small positive effect of military expenditure on own-country growth. They conclude that "the link between military spending and economic growth is complemented by a strong economic spill-in from neigboring states" (p. 812).

This approach is updated and extended by Yildirim and Öcal (2014). They add an augmented Solow growth model for which they find a suitable econometric specification using a somewhat modern technique, due to Hoover and Perez (1999).[6] Despite the different time period studied by these later authors, and the completely different method of arriving at a model, their conclusions are identical to those of Shin and Ward, that there is "a synergy between military expenditure and economic growth" (p. 101) and that neighbour-effects are strong for both milexp and economic growth.

A genuinely global approach is provided by Nordhaus, Oneal, and Russett (2012), although their focus is the precise opposite of the other studies mentioned here. They attempt to find the effect of the international security environment on milexp across 165 countries. To do this they develop an index which is in some way related to "the probability of becoming involved in a fatal militarized interstate dispute."[7] This index involves each country's autocracy–democracy score, the trade interdependence of the pair, GDP, the geographical distance between them and the number of year of peace since the last conflict between them. Using a binary variable of 'did' or 'did not' have a conflict in a given year, they estimate the coefficients of each of these covariates, and are then able to predict a probability of conflict from any given set of covariates. The total probability of a conflict in a given country is then simply the combination of these (presumed independent) probabilities.

This conflict probability is then used as a covariate in a panel data model with military expenditure as the dependent variable and many other covariates included. They find that this constructed probability of conflict has more influence on milex the milex of friends or foes, as well as the actual incidence of war or the severity of that war. This is an impressive study, and one which Baudains returns to in his work on threat which is central to the work of this chapter.

---

6 The technique of Hoover and Perez (1999) is known as general-to-specific and is a simple procedure for finding an appropriate (somehow-defined, usually via goodness-of-fit measures or information criteria) empirical model. One begins by including every covariate available, then run a series of regressions in which variables are removed until such time as the chosen metric is deemed to be optimised.

7 The term fatal militarized dispute is defined as one in which at least one combatant dies.

9.2.4 Measures of threat

We now turn our attention to the spatial measure of threat developed by Baudains, Fry, et al. (2016) and Baudains and Wilson (2016) which will be central to our attempt at a truly global model of military expenditure.

Attempts at measuring threats between nations can be traced back to the arms-race work of Lewis Fry Richardson who investigated the influence of adversary spending abroad in increasing military spending at home using a simple mathematical model (Richardson 1919; Richardson 1960). Annual military spending, p and q, in two rival countries is modelled using coupled differential equations of the form:

$$\dot{p} = -\sigma_1 p + \rho_1 q + \varepsilon_1$$
$$\dot{q} = \rho_2 p - \sigma_2 q + \varepsilon_2$$

Richardson then extends this to model N interacting nations in an obvious, additive manner adding a ρ parameter for each country pair. But, as Baudains (2016) points out, this model lacks a broader international security environment, since it is based only on country pairs.

The complexity of global decision-making when looking at threats and alliances with all countries in the world is described vividly by Baudains:

> when considering the threats arising from a nation j, nation i will not only consider its relationship with j and adjust its spending accordingly, but also look at the international security environment... If j increases their expenditure rapidly, and i is an historic adversary of j, then i will be likely [follow suit]. However, if j is known to [be an adversary of] k, then i might decide they do not need to adjust their spending by quite so much, since k is likely to [do so] on behalf of i (particularly if k is in an alliance with i), negating the threat posed by j. (Baudains 2016, p. 205)

Baudains therefore takes Richardson's very simple model, letting p and q stand for some broad measure of hostility, and allows it to evolve over a number of locations in space, with the p at each location dependent on the q at every other location. Weighting factors are introduced on each q to allow for both distance between the countries and a measure of "dyadic alliance" between the two countries. This alliance measure allows countries to respond differently to one another's aggression, depending on any historic alliances between them.

This alliance measure is constructed inspired by the work of Nordhaus, Oneal, and Russett (2012) to use the measure developed by Signorino and Ritter (1999) of the similarity of foreign policy positions as a measure of friendship or otherwise. This latter work proceeds from a Correlates of War alliance measure with 0=no alliance, 1=entente, 2=neutrality or non-aggression pact, 3=mutual defence pact.[8] Pairs of countries are compared for the number of instances where they have the same value of alliance with each other country. This gives each country a simple alliance-similarity measure. They combine these with other dyadic similarity variables (including, but not limited to, distance) into a linear combination where some variables are weighted more highly than others.[9] This weighting allows, for example, an alliance with a militarily 'important' state to count more than one with an unimportant state.

Despite these subtleties, and indeed the work of Signorino and Ritter (1999) allows for a great deal of nuance and expert input, Baudains uses the simplest possible version of their alliance measure. For Baudains, all weights are equal (implying that each alliance is as important as each other) and the only other correlate is the geographic distance between them. Baudains acknowledges that this is merely a "proof of concept" (p. 213), but the naivety of the measure produces some unusual results in what follows, and we will point out some oddities in the results in the following section. We then describe in section 9.4 on p. 245 how we combine the demonstration model and Baudains' measure to explore the global implications of military expenditure.

[8] Correlates of War is an almost ubiquitous international security database available at correlatesofwar.org, and the alliance data is due to Gibler (2009). Signorino and Ritter further allow that states have mutual defence pacts with themselves, for completeness.

[9] Another quite interesting example of a dyadic relationship between states is the extent to which their "UN vote portfolios" match, i.e. how often they have voted in the same way at UN votes.

9.3 Discussion of Baudains' threat measure

As discussed earlier, Baudains dyadic threat measure is based on the sound and subtle theoretical work of Signorino and Ritter (1999). Unfortunately, Baudains' application of this work represents a "simplest possible" approach to operationalising their theory, and hence has some entries which are not in line with what one might expect from a general knowledge of the global security environment.

Many of the most unexpected features of Baudains' spatial measure of threat stem from apparent irregularities in the alliance similarity measure which underlies the threat measure. This may be compounded by the equal weighting he uses, as discussed in section 9.2.4. This allows all alliances equal influence on the final measure, an alliance with Haiti counting equally as an alliance with the USA. The second cause of issues with the estimates is the relative balance between alliance similarity and distance in Baudains' measure. Some examples of irregularities in the alliance similarity measure follow.

For the USA, their most allied country is Canada. This is a satisfying place to start and seems intuitively correct since it is nearby and has NATO and NAFTA trade relationships. But following Canada is a number of Caribbean Islands, such as Haiti and Jamaica. These come before Great Britain which is eight in the US's list of allies. Some of the USA's allegiance similarities are harder to explain still: Both Bolivia and Brazil are considered a closer ally of the USA than Norway. The bottom end of the USA's alliance table also has some puzzling features. For example, Saudi Arabia is around 10 places lower than Russia, and bottom-place Mauritania has a negative alliance score (albeit very close to zero). It is not clear how this can be arrived at mathematically, and suggests a possible error in the calculation.

The overemphasis of the distance measure has a less noticeable effect for the UK, because European partners are plausibly highly represented. The first six most closely-aligned are the Netherlands, Belgium, Luxembourg, Spain, France and Portugal. Mauritania also occupies bottom position for the UK, but its alliance measure is far higher (0.4) than its alliance score with the USA (0.0, with 1.0 being the theoretical maximum). The USA though is a long way down the list of UK allies, slightly below Mali and Tunisia.

More concerning still is the fact that the alliance similarity measure has unusual patterns which are must be an artefact of either the method used or the data which calibrates the method. Iran and Israel have identical levels of alliance similarity with every country other than themselves (since every country has an alliance similarity of unity with itself). The leads to the absurd situation that the USA has identical alliance similarity (0.32) with Israel as it does with Iran. And 46 countries have a numerically identical alliance with the UK, suggesting the alliance measure may be too blunt an instrument to distinguish between, for example, South Korea and Zimbabwe.

The inclusion of military expenditure in the threat measure which is, in part, based on this flawed alliance measure, results in further unexpected results. The USA has by far the largest military expenditure of any country and this leads it to represent a high threat to countries which one might not expect. Encouragingly, Russia and Cuba are both in the top 3.[10]

But Japan is surprisingly high (11th place) as are Ireland (13th) and Iceland (14th). The other end of the scale is more encouraging, with the UK, Canada, Israel and Belgium all featuring appropriately low down the list. The countries which threaten the USA seem to be countries with high military spending, with China, Russia and India taking the top three places. But France, Great Britain and Japan are all also within the top 10. This suggests that the measure overstates military spending and understates the alliance measure.

Threats to the UK are more problematic again. The three biggest threats to the UK are the USA, France and Italy. Once again, we see the overweighting of military spending versus the alliance measure which, at least in the case of France and Italy, represents reality more closely.

Overall, as we have seen, the results for both the alliance similarity and the threat measures are not at all encouraging. We therefore offer what follows purely in the spirit of demonstrating the kind of analysis and visualisation that might be undertaken using such a dyadic threat measure, were one to be developed with properties which pass basic scrutiny.

[10] Mauritania receives the highest threat from the USA, greater than Russia by a factor of 2. This is an irregularity no doubt caused by the irregularity with the alliance measure mentioned above.

9.4 Modelling responses to threat

Given a measure describing how threatened a country feels, we can imagine some commensurate increase in spending on military hardware and personnel the effects of which can be modelled as an economic and trade phenomenon using our model. By making some assumptions about the nature of this additional spending, we can bring our model's two main metrics (GDP and balance of trade, see chapter 3 for full details) to bear and move from a discussion of how much military expenditure is being generated by a particular threat environment to one of which countries stand to gain and lose the most economically from those expenditures.

Unlike the migration we studied in chapter 8, international security is not described naturally in terms of dyadic relationships.[11] As we have seen, the work of Baudains, Fry, et al. (2016) has produced a global spatial model which allows for the calculation of a set of threat dyads, T_{ij}, which combine the level of military spending of countries i and j with a continuous measure of the extent of the military alliance between them. Thus, countries with high levels of military expenditure and low levels of alliance will tend to have high levels of threat between them. Geographical proximity additionally tends to aggravate this threat.

This threat measure is shown to be a significant predictor of military spending, with a 1% increase in threat leading to a 0.06% increase in the military expenditure of the threatened country. This relationship gives us the link between the fields of international security and economics which we have been looking for.

The question remains, though, of how to model military spending in the model we have developed. WIOD has no sector related explicitly to such expenditure. (And, indeed, the data which would be needed to capture the demand for such a sector would most likely be considered too sensitive for publication.) We thus make the simplifying assumption that military expenditure comes purely in the form of final demand for the output of the metals sector.

This is clearly a gross simplification. And it is one which will colour the exploratory results of this chapter significantly. It would clearly be an interesting avenue for future research to relax this assumption by using a typical "basket" of sectors which make up the average unit of military expenditure. But our purposes here are not to design the perfect economic measure of military expenditure, but to show how we might gain

[11] By "dyadic" we mean having two 'endpoints'; one source and one destination in the case of directed dyadic relationships. In our case these endpoints are countries.

global context in the realm of international security by attaching it to an economic model which will allow for some interesting global results.

9.5 Alliance similarity and trade

We now know how an increase in dyadic threat affects the coefficients of the model by increasing demand for the metals sector. It seems natural that in fulfilling the part of this particular increased demand which must come from imports, countries are more likely to turn to their allies than to their foes.

In order to model this dynamic we need a measure of which countries a particular country feels most and least allied to. For this we will use the measure of "alliance similarity" given by Baudains (2016) and discussed in detail in section 9.2.4 on p. 241. As we discussed there, this measure works by comparing the alliance membership portfolios (e.g. membership or not of NATO etc.) of each pair of countries and assigning a score based on their similarity. Thus, two countries are considered to be close allies if they have a large number of allies in common.

We extend the interpretation of alliance similarity, s_{ij}, to include the idea that closer allies are more likely to trade the metals which are earmarked for military use. Existing patterns of metals trade, y_{ij}, should also influence trade increases for military use. The simplest way to combine these two measures is to simply multiply them, creating a new 'security-related trading preference' variable $s_{ij}y_{ij}$. This has the advantage that it does not require a new parameter to distinguish between the relative strength of these effects, as would the standard economics way of combining effects via a Cobb-Douglas or CES functional form. A more in-depth exploration of this topic would involve estimating empirically such parameters but this is beyond the scope of this demonstration study.

We can abstract away from much of the problem of interpreting our new trading preference variable by scaling it as a fraction of the variable's total value across all trading partners. In this way, it becomes a unitless measure of how much a particular country is preferred as security-related trading partner compared to other countries. In a system with C countries in total, countries with a scaled trading preference variable greater than '1/C' will be favoured for security-related trading to a greater-than-average extent.

We therefore have a very simple expression relating an increase in country j's security-related demand $f_{j,\text{metals}}$ to the related increase in a particular trade flow, $y_{ij,\text{metals}}$ and the scaled security-related trade preference measure:

$$\Delta y_{ij,\text{metals}} = \Delta f_{j,\text{metals}} \frac{s_{ij} y_{ij,\text{metals}}}{\sum_k s_{kj} y_{kj,\text{metals}}} \tag{9.1}$$

Each flow change, $\Delta y_{ij,\text{metals}}$, is the total new trade between i and j due to an increase in military spending in j.

In summary, this model hypothesises that the suitability of the trading partner in the face of an uncertain security environment depends on both the level of current trade and their alliance structure. By combining this with our assumption that an increase in threat of 1% increases metals expenditure by 0.06%, we can now model how the global metals trade would respond to a particular increased threat.

9.6 Results

For the purposes of this experiment we will restrict ourselves to studying only the largest portion of the dyadic threat measure. We restrict our analysis to the most important threat relations, $T_{ij} \geq 1$, which leaves us with 399 dyadic relationships of a possible 16,770 in the dataset.[12]

9.6.1 Threat summary statistics

Table 9.1 shows summary statistics for the threat measure, specifically the ten largest sources and targets of threat among our sample of 399. As with all of the work of this section, we present these analyses here merely as an indication of the kind of analysis which could be done with a more robust threat measure. See section 9.3 on p. 243 for more on how and why this threat measure is flawed.

By far the most common source of threat comes from the USA which is the source of 127 of the 399 largest threats. The UK is a distant second with 61 followed by France (55), India and Saudia Arabia (32), and Italy (27). The USA is also the most common target for threat, although the targets are far more evenly distributed than the sources. Canada, Iraq and Libya follow closely behind. This preponderance of the USA is an artefact of the fact that the threat measure includes military spending, and US military spending is around triple that of China's, with the second biggest expenditure.

A visual overview of the entire network of 399 relationships is shown in figure 9.1 in the form of a weighted, directed graph, with countries represented as nodes, and the level of threat between the nodes as edges of varying weight. The nodes are sized according to the amount of threat they emanate (or their "weighted out-degree" to use the graph theory terminology.) Additionally, nodes have been coloured according to the results of a community detection algorithm which was configured to find precisely three communities. Each edge is then coloured with a linear combination of the colours of its two end nodes.

The clusters, which are visually quite distinct, at first appear surprising. If this is a network of threat, not alliance, why are the large (i.e. threat-producing) European countries in a single cluster (orange, bottom left)? Closer inspection reveals that the large European nodes are not in a cluster because they threaten *one another*, rather

[12] Dyads involving the neutral countries Costa Rica, San Marino, Switzerland and Turkmenistan are removed from the analysis. These nations were susceptible to generating low levels of alliance measures since they do not belong to many of the agreements that would otherwise be captured by this measure (e.g. Switzerland does not belong to the EU or to NATO). They are also unlikely to be the source or targets of threat as a result of their neutrality.

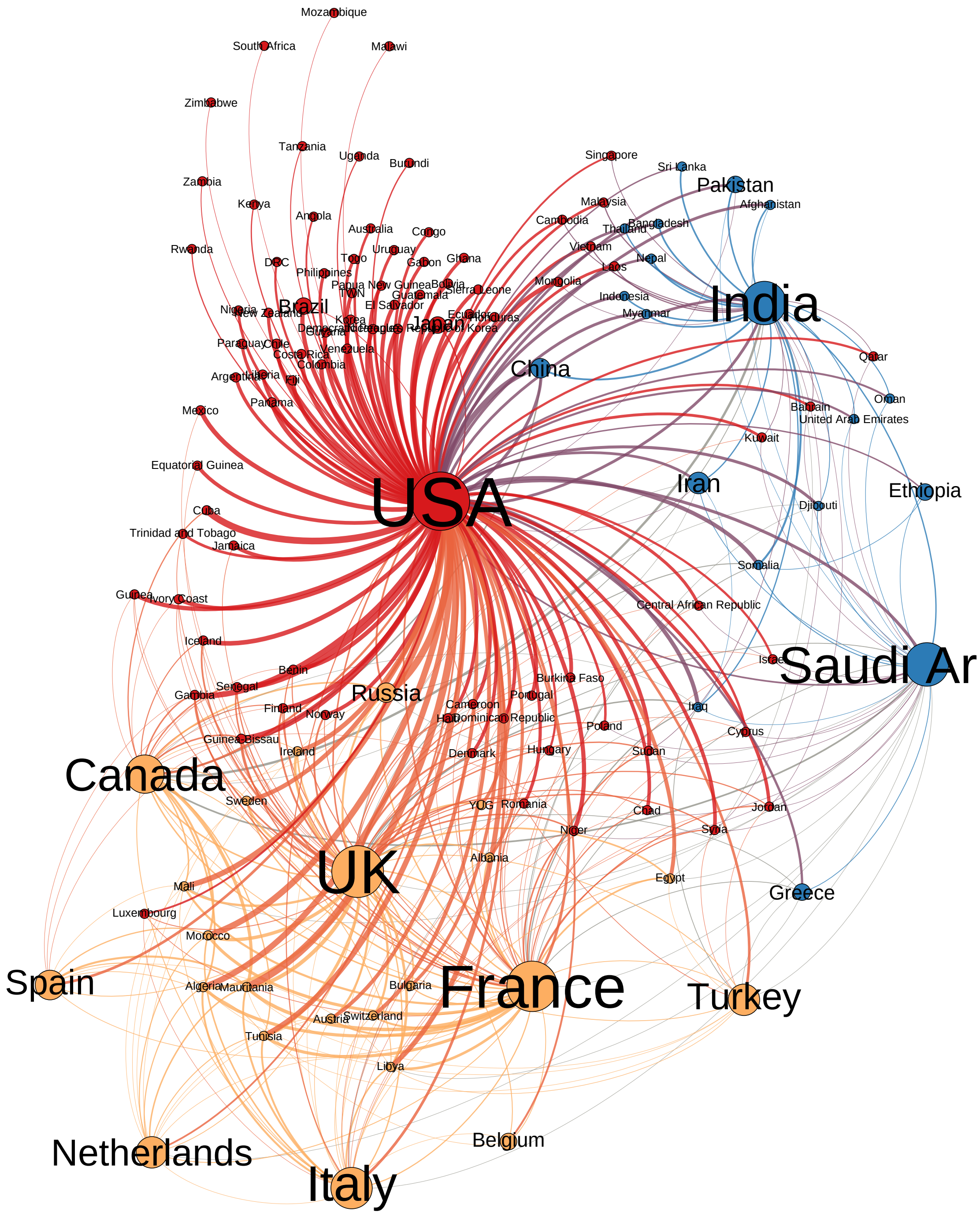


Figure 9.1: The 399 largest dyadic threat relationships (those with $T_{ij} \geq 1$). Node size is proportional to the weighted out-degree, meaning that larger nodes create more threat. The thickness of the edges is proportional to the logarithm of T_{ij}. Nodes are grouped into three clusters using a community detection algorithm.

Table 9.1: The ten most commonly occurring sources and targets of threat among the subset of $T_{ij} \geq 1$. Refer to section 9.3 on p. 243 for a discussion of the oddities in these estimates.

(a) Threat sources

Country	Instances
USA	127
UK	61
France	55
India	32
Saudi Arabia	32
Italy	27
Canada	21
Netherlands	11
Turkey	11
Spain	9

(b) Threat targets

Country	Instances
USA	12
Canada	10
Iraq	9
Libya	9
Mauritania	9
Tunisia	8
Mali	8
Russia	8
Algeria	8
Egypt	7

because they threaten the same set of other countries. Notably this group includes the many of the countries of North Africa: Morocco, Algeria, Tunisia, Libya.

The USA is in a cluster all of its own, as a threat producer. Pinned between the European threat-producers and the USA are Russia and many sub-Saharan African countries.

The final cluster could perhaps be described as “other”, consisting of the large threat producers which are not European and not the US. China, India, Saudi Arabia, Iran, Pakistan and Ethiopia are the main threat producers in this cluster. The South-East Asian countries of Singapore, Malaysia, Cambodia, Vietnam and Laos are all sandwiched between the USA and this third cluster of threat-producers. Finally, the Middle Eastern countries of Jordan Syria and Egypt appear visually to be threatened almost equally by all three.

9.6.2 Results of the experiment

One simple way of assessing the effect of dyadic threat is to measure which third-party country stands to benefit by some metric when that threat is increased.[13] This benefit will be brought about by an increased production of the metals sector, due to increased demand from a threatened country which is either a large importer of metals or a close ally, according to equation (9.1). The additional production will, depending on the amount of value added per unit of the metals sector in that third party country, have an affect on total value added (GDP).

[13] We use the term “third-party” in the sense of being not the source or the target of the increased threat.

By applying equation (9.1) to each threat dyad in our sample one at a time, and measuring the results, we have a simple experiment which will show us how the international security situation might affect the global economy. In each case, the threat level is increased by 1%.

Table 9.2 shows the third-party countries which are most affected by the threat dyads in our sample, first by total benefit summed across all 399 dyadic threat increases in (a) and then by total benefit as a percentage of GDP in (b). Notice that we are excluding the possibility of benefits to either the threat-producing or the threatened country, although there will clearly be effects in both.

The two tables are markedly different. By total benefit, the top 7 are among the world's largest economies: Canada, Italy, China, France, Russia, Japan and the UK. This reflects the fact the large countries are more likely to be substantial existing exporters of metals and therefore have the most to gain from a third-party increase in demand. But this is not simply a measure of economy size: the USA, Brazil and South Korea all come much further down the list. In the case of the USA, this is presumably because it is involved in so many of the 399 dyadic threat relationships and therefore cannot count as a beneficiary of any increase in demand as a result.

The list by percentage of GDP is, perhaps predictably, dominated by smaller countries, although several with a GDP of over $10 billion make it onto the list, including Canada, South Africa, Venezuela, Switzerland and Norway.

The top three are all large metals exporters. The West-African nation of Guinea has a large aluminium ore industry, which accounted for $600 million of exports in 2005, over a third of the global trade in this product (Simoes and Hidalgo 2011). In the same year, Zambia exported almost $3.8 billion of copper, 72% of that country's total exports. Finally, Papua New Guinea's large exports of copper (3% of global exports) and gold (around 35% of its total exports) may account for its presence at the top of the table.

Clearly gold is not a metal primarily associated with military expenditure. This is an artefact of the simplicity of the approach we are taking in this chapter. Our suggestion of future work involving a basket of consumption goods associated with military expenditure will help here, although no simple solution will allow us to disaggregate gold and precious metals from the non-precious industrial metals which are combined into one sector in WIOD.

Despite these caveats, Guinea taking first place in this table is still a surprising result and deserves a little more inspection. In order to examine in more detail how Guinea

Table 9.2: The largest total beneficiaries of increased threat across the 399 dyadic threat relationships under experiment. Only third-party benefits are considered. Benefits are summed across all threat increases.

(a) The top 20 by total benefit.

Country	Benefit ($ M)	GDP ($ Bn)	Benefit (% GDP)
Canada	22.45	914	0.0025
Italy	7.93	1561	0.0005
China	7.43	1511	0.0005
France	6.69	1886	0.0004
Japan	6.23	3600	0.0002
Russia	6.16	671	0.0009
UK	5.91	2162	0.0003
Mexico	4.73	642	0.0007
Belgium	4.00	415	0.0010
Netherlands	3.86	676	0.0006
Spain	3.79	877	0.0004
Brazil	3.34	725	0.0005
Switzerland	3.28	295	0.0011
Korea	3.13	606	0.0005
South Africa	3.13	186	0.0017
Sweden	3.01	361	0.0008
Australia	2.94	563	0.0005
Venezuela	2.60	164	0.0016
USA	2.48	11230	0.0000
Norway	2.41	225	0.0011

(b) The top 20 by percentage of GDP.

Country	Benefit ($ M)	GDP ($ Bn)	Benefit (% GDP)
Guinea	0.22	3	0.0073
Papua New Guinea	0.23	4	0.0054
Zambia	0.31	6	0.0051
Mozambique	0.24	6	0.0041
Chile	2.36	95	0.0025
Canada	22.45	914	0.0025
South Africa	3.13	186	0.0017
Venezuela	2.60	164	0.0016
Guyana	0.01	1	0.0014
Brunei Darussalam	0.13	10	0.0013
Malaysia	0.95	75	0.0013
Congo	0.11	8	0.0012
Equatorial Guinea	0.12	11	0.0012
Oman	0.31	27	0.0011
Switzerland	3.28	295	0.0011
Luxembourg	0.61	56	0.0011
Yemen	0.14	13	0.0011
Jamaica	0.10	9	0.0011
Norway	2.41	225	0.0011
Libya	0.41	39	0.0011

Table 9.3: The threat dyads causing the biggest increase in the GDP of Guinea, following a fixed percentage increase in the threat level. The 10 largest effects are shown. The existence of threats between the USA and both France and the UK which are greater than that with China is a reminder that Baudains' threat measure overstates the importance of military expenditure over alliance similarity.

Threat from	Threat to	ΔGDP_{Guinea} ($ '000s)
USA	Russia	191
USA	France	7.1
USA	UK	5.7
China	USA	2.9
India	USA	1.5
USA	Korea	1.4
USA	Spain	1.1
USA	Italy	0.8
USA	China	0.7
USA	Ireland	0.6

benefits from threat increases, we look at which of the 399 dyadic threats served most to increase Guinea's GDP. The 10 largest dyads are shown in table 9.3.[14] As with the overall dyadic threat dataset, threats from the USA dominate the list, with that between the USA and Russia coming at the top by a very wide margin. Almost 100% of Guinea's exports to Russia in 2005 were aluminium ore (Simoes and Hidalgo 2011) and although this flow accounted for only 0.2% of Guinea's total exports of this product, it represents 10% of all Russia's aluminium ore imports.

[14] A more complete picture of the dyads which increase Guinea's GDP is shown in figure 9.2 which is a network representation of the dyads, with edge weight being proportional to the increase in Guinea's GDP induced by the increase in the dyadic threat. The fifty largest dyads are shown.

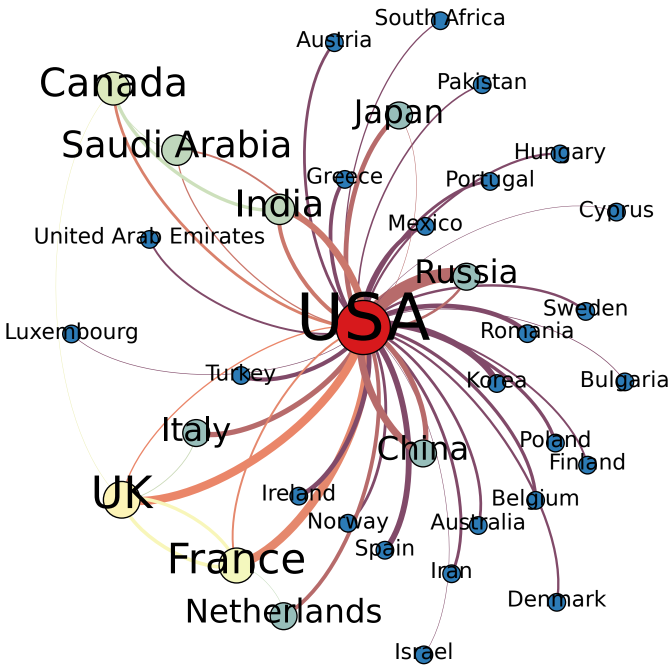


Figure 9.2: A network diagram of the 50 dyadic threat relationships to which the GDP of Guinea is most sensitive. Node size (and colour) is proportional to node degree. The weight of the edges is proportional to the increase in Guinea's GDP following a fixed-percentage increase in T_{ij}, the dyadic threat measure.

9.7 Conclusion

In a similar manner to that presented in chapter 8, we have presented an extension to our model involving the addition of information derived outside the model. But whereas we treated the migration flows in that earlier analysis as coming from "data", here the threat dyads we augmented our model with were more explicitly the result of a separate modelling process. We draw this distinction here to demonstrate how our model can be useful in either setting.

Using the work of Baudains, Fry, et al. (2016), however flawed we know that work to be, we developed a very simple model of how an economy responds to threat. Namely that it increases military spending by 0.06% and that that spending occurs entirely within the metals sector. This allowed us to study the effect of an increased threat between any given pair of countries on the rest of the countries in the model.

We found that in percentage of GDP terms, the country to most benefit in total from a 1% increase in each of the threat dyads we explored was the West-African nation of Guinea. Guinea was shown to be a major exporter of aluminium ore to Russia.

As discussed earlier, the considerable problems with the threat dyads used mean that these results cannot be taken at face value, but rather as a demonstration of the kinds of analysis, and visualisation of results, which could be carried out with a more reasonable measure of threat. But we nevertheless begin to see how the additional context of the global trade network might result in interesting corollaries of the findings of other social science models, including how Russia and the US increasing their military expenditure may have economic benefits and costs in unexpected parts of the world.

Chapter 10

Trade facilitation and the global economy*

10.1 Introduction

We have now arrived at the last of our extensions of the model we have developed. As with the extensions of chapter 8 and chapter 9 our aim will be to show how the inclusion of a global economic context, allows modellers to explore a social science question in broader, more international terms. This chapter's extension concerns export infrastructure and the relationship between trade facilitation and macroeconomics, particularly in an economic development context.

A regular contributor to the literature on trade in the least developed countries (LDCs), John S. Wilson, has defined trade facilitation as follows:

> Trade facilitation most often implies improving efficiency in administration and procedures, along with improving logistics at ports and customs. A broader definition includes streamlining regulatory environments, deepening harmonization of standards, and conforming to international regulations. (Wilson, Mann, and Otsuki 2003, p. 367)

According to the World Trade Organization (WTO), many developing countries face "a range of supply-side and trade-related infrastructure obstacles" which limit their ability to participate in global trade.[1] Aid for Trade is a WTO-led initiative to encourage aid donors to use aid money to alleviate these trade-inhibiting difficulties.

---

* This chapter uses the 200-country model throughout. See page 77 for a description of what this means.

[1] These quotes are taken from the WTO's "Aid for Trade" page under the rubrics of "trade topics" and "development," available at bit.ly/1zGS6R1.

Many existing works on trade facilitation are country-specific case studies looking at trade reforms, such as Soloaga, Wilson, and Mejia (2006), or econometric studies looking at the effect of aid on increasing trade volume rather than on the more pertinent issue of GDP growth, which we are able to study here.[2] A notable example by the World Bank itself, (Helble, Mann, and Wilson 2012), finds that Aid for Trade boosted exports by $1.33 for every $1 of aid given. These latter regression analyses do not attempt to assess the global effect of an increase in trade in one country on the global trade network in general. They also focus on the outcome of increased trade. We can go one step further here and focus on GDP growth, which is the eventual goal of increased trade.

We create a very simple trade model based on the type of spatial interaction model outlined by Wilson (1971). We assume that trade facilitation policies increase the attractiveness of a particular sector in a recipient country as a trade partner, thus tending to increase the volume of goods which importing countries choose to purchase. The impact of this increased exporting by the recipient country is then assessed in GDP terms using the demonstration model.

A reduced export price in one country has two effects, a wealth effect and a substitution effect. The wealth effect tends to increase demand, and the substitution effect tends to replace other sources of the good with the now cheaper export in question. Which effect dominates is not *a priori* certain, but it is clear that a decrease in export price in one nation can only hurt competitor nations, even if the wealth effect allows for more of a good to be import overall than before. Although this kind of general-equilibrium price elasticity is absent from the demonstration model, with some simple assumptions about the elasticity of substitution between products from different countries, we can have a first attempt at assessing the impact of trade facilitation both in terms of the benefit to the recipient country, and of the loss to existing exporters.

Before turning to this modelling exercise, we first review the literature on trade facilitation, looking first at the majority of studies which use gravity models, then at the few examples to use CGE modelling, particularly the GTAP system. We finally give a brief overview of some of the specifics of trade facilitation, such as administrative improvements and reduced transport costs.

[2] A good review of this literature is given by Suwa-Eisenmann and Verdier (2007)

10.2 Literature Review

John S. Wilson, whose work appears throughout the literature in the 2000s, described the empirical literature on trade facilitation in 2003 as "limited" (Wilson, Mann, and Otsuki 2003). Before 2000, much of the literature on LDCs and trade was focused on liberalisation policies. It was particularly focused on proving, or otherwise, that LDCs which removed trade restrictions received an attributable benefit to growth. This literature is extremely voluminous and is not of direct interest to us here, but Chicago-school economist Sebastian Edwards gives a representative example of the pro-free trade approach of the 1990s in building a model in which trade openness is assumed to lead to increased innovation and hence that "countries that liberalize their international trade and become more open will tend to grow faster." (Edwards 1992, p. 31)

It is not until the early 2000s that the focus turns to enable the export effectiveness of developing countries, with studies on how this might be done and modelling the possible effects. The vast majority of these use a gravity model, with the exception of a few studies which use CGE models (often GTAP). In this short literature review, we will survey the most well-known gravity models of export effectiveness in LDCs, then turn our attention to CGE/GTAP models. We will finally give a brief list of the studies which look at particular elements of export effectiveness such as infrastructure and administrative efficiency.

10.2.1 Gravity models of trade facilitation

Perhaps the earliest use of gravity models to study trade in LDCs, beyond the 1980s and 1990s debates around trade openness, is the study by Otsuki, Wilson, and Sewadeh (2001) of how food safety standards in the EU affect exports from African countries. Although their approach is described by Wilson, Mann, and Otsuki (2003) as being "somewhat crude" it is commended for being "less data demanding and more applicable for developing economies whose price data are less reliable and complete" (p. 374).

They are interested in modelling the effect of a change in regulation around a certain toxin which contaminates some foods. This is not of direct interest to us here, but their approach is a useful starting point. Their gravity model estimates the *value* (not quantity) of trade, disaggregated by product category, k, from each African country, j, to each EU country, i. Preceding the more advanced theoretical work which would shortly follow, their gravity model is very much in the tradition of early practitioners such as Tinbergen and Pöyhönen (see chapter 2). They express their formulation in logs as:

$$\ln V_{ij}^k = b_0^k + \sum_m b_m^k \ln X_{mij} + \varepsilon_{ij}^k$$

The X variables are per-capita GDP for both countries, the distance between them, a year dummy, a colonial ties dummy and their toxin regulation variable.

With every variable being dependent on k, their gravity model is really a set of unrelated gravity models, one for each product category. This implies that trade in one good is unrelated to trade in another. Furthermore, it doesn't allow for an increase in trade from one country to directly affect trade in another country. The work of Anderson and Wincoop (2003) (hereafter AvW) on "multilateral resistance", described in some detail in chapter 2, in the section beginning on p. 29, undid these separations between the equations, and brought a level of theoretical 'gravitas' to the gravity model. It resulted in an explosion of gravity studies, including those relevant to us here.

The first paper in this literature to acknowledge the existence of this new work was that of Wilson, Mann, and Otsuki (2003). Although it is mentioned by them in their introduction they do not use it, favouring instead the simpler approach of Otsuki, Wilson, and Sewadeh (2001), saying that their approach was "designed to overcome limited data availability" (p. 379).

They study the effect of four aspects of trade facilitation on the global trade patterns of countries in the Asia Pacific Economic Cooperation group (APEC), each one taken from particular survey data for each of the countries.[3] Their four aspects are port infrastructure; the customs environment (customs costs and "administrative transparency); the regulatory environment; and "e-Business usage," a measure of the extent to which telecoms, financial services and logistics are able to use networked information to improve efficiency. To Otsuki's formulation, they also add the GNP of each country (alongside the GDP per capita which Otsuki also used) and a bilateral fractional tariff term along with their four export effectiveness terms. Finally, they also include exporter fixed effects and time fixed effects.

Estimating their model seemingly with ordinary least squares, although this is not specified, they find find that port efficiency is the most important of their four factors, seemingly able to make a strong positive effect on trade flows.[4] They also find that the

[3] John S. Wilson's centrality to this literature, and his particular interest in APEC, has led to the group having an odd over-representation in the literature around trade facilitation. But he has also written on Mexico and Sub-Saharan Africa.

[4] In contrast to other quantitative economics papers in non-development contexts, Wilson, Mann, and Otsuki (2003) are typical in this literature in paying little attention to the theoretical justification of their statistics work. Many of their point estimates are significant at 1% (in fact, in their preferred

regulatory environment has a negative effect on trade flows, concluding that "tightening regulations can offset improvements in other trade facilitation measures" (p. 379).

Work with gravity models which do not take the work of AvW into consideration continued with, for example, Linders et al. (2005), who use a Tinbergen/Pöyhönen model to study the effect of institutional quality and "institutional distance" on trade values. The former is an average of six governance variables and the latter is proxied for by a measure of "cultural distance" which involves things like shared language and shared religion. They find that institutional distance has the expected negative effect on trade but that cultural distance has a positive effect. They explain this with an appeal to the international business literature: "Because of the high costs and uncertainty of successfully operating production facilities in culturally distant countries, firms expanding into such countries tend to opt for entry modes requiring relatively little resources, such as exporting." (Linders et al. 2005, p. 17)

Attempts in this literature to take seriously the work of AvW were slow to come. The study of Mexican trade competitiveness of Soloaga, Wilson, and Mejia (2006) comes close, describing the work of AvW in some detail, but do not implement it, leaving that task to "a further version of this paper" (p. 9). They use an identical set of indicators of export effectiveness as Wilson, Mann, and Otsuki (2003), now claiming that the e-commerce indicator is a proxy for service sector infrastructure. They additionally update that earlier work to use an estimation procedure suggested by Silva and Tenreyro (2006) who had recently found theoretical difficulties with estimating gravity models using OLS, and proposed an estimator which could account for zeros in the flow data. Soloaga et al. find again that port efficiency has a positive impact on trade, but this time finding that the regulatory environment has a positive impact too. This is now justified by saying that it proxies for the perception of corruption.

In a paper which focuses on Africa, Iwanow and Kirkpatrick (2009) use indicators of trade facilitation, quality of regulation and infrastructure to study how these things contribute to export effectiveness. This is the first paper in the literature to use the two-step procedure of Helpman, Melitz, and Rubinstein (2008) which first estimates which countries trade with one another *at all* and then, in the second step, estimates

specification, every point estimate is significant at 1% except for an adjacency dummy which is not significant, and the customs environment variable which is significant at 5%). This looks certain to be related to the fact that their "sample size" is large: they report $N = 3,304$. But this is panel data: each observation is for a particular country pair in a particular time period. It is questionable whether these observations can be treated as random draws from a large population, as OLS estimation requires. There are standard methods of dealing with this problem of course, but Wilson et al. make no mention of such econometric specification questions. Silva and Tenreyro (2006) find major theoretical problems with estimating gravity models using OLS, and their findings are refined by a later Monte Carlo study by Martínez-Zarzoso (2013).

the gravity model itself. For a fuller discussion of the well-known work of Helpman et al. see chapter 2 and the section beginning on p. 29. The authors briefly mention AvW, and have an innovative method of accounting for the multilateral resistance term. They include a "remoteness" variable which is a GDP-weighted average distance to all other countries. This is a neat mathematical trick, but does not allow for the interlinking of unrelated countries' trade flows which AvW envisaged.

Further similar gravity models are estimated by Korinek and Sourdin (2009) who use a similar specification to estimate the effect of maritime transport costs on trade using a then newly-compiled OECD database; Hoekman and Nicita (2011) who look at trade policies, measures of trade facilitation and a "Logistics Performance Index" (LPI), both from the World Bank, for LDCs; and Portugal-Perez and Wilson (2012) who look in detail at the various trade facilitation measures, separating them into "hard" and "soft" infrastructure, the former being tangible and the latter training or efficiency-related.

Fully "modernised" gravity models, which include AvW's multilateral resistances, Helpman's two-stage selection-into-trade model, and the Taylor-approximation solution method described briefly in chapter 2, first appear with Behar, Manners, and Nelson (2013) and Francois and Manchin (2013).

Behar, Manners, and Nelson (2013) use the LPI to produce a very theoretically-complete gravity model which they estimate from 88 low- and middle-income exporters and find a positive and highly significant role for their exporters' logistics index in explaining trade flow values. A similarly rigorous approach is taken by Francois and Manchin (2013) who use an institutional-quality index from the "Economic Freedom of the World" database by the Fraser Institute, and a series of infrastructure measures, such as percentage paved roads and number of mobile phones per capita from the World Development Indicators database of the World Bank. They find that not only are poor infrastructure and poor institutions a significant hindrance to LDC exports, but also that they "limit market access for exports from the North" (p. 173).

10.2.2 CGE and GTAP models of trade facilitation

As with the literature using gravity models, the LDC trade literature using computable general equilibrium (CGE) from the 1980s and 90s is overwhelmingly focused on the effects of trade liberalisation, inspired perhaps in part by the 1996 'Manila Action Plan for APEC' which called for the reduction of tariff and non-tariff barriers throughout the APEC region.[5] The CGE modelling of trade facilitation was still a rare enough

[5] For details of CGE modelling, and the GTAP database/model in particular, see the section of the literature review beginning on p. 56.

undertaking in 2009, that one researcher was moved to claim that "trade facilitation has never been explicitly modeled in CGE models" (Zaki 2009, p. 3). As we shall see this is not quite the case, the OECD mention a few such studies (Walkenhorst and Yasui 2003), but liberalisation did continue to be the most often-studied phenomenon in CGE modelling of LDCs.

In studying a free-trade agreement between Japan and Singapore, Hertel, Walmsley, and Itakura (2001) model trade facilitation via price reductions. These include the direct costs reductions from customs automation, which the authors estimate through original research with the Japanese government itself, and through an earlier estimate by GTAP's David Hummels that a one day delay in trade completion is equivalent to a 0.5% *ad valorem* extra cost.[6] By estimating the number of days saved by the new trade deal, they are able to convert time costs directly into price reductions.

This then allows the authors to exploit the considerable power of the global GTAP database to make some forecasts about the effects of the trade deal which encompass the wider context. They give predicted effects for both Japan and Singapore, as well as for the rest of the world, finding that GDP is boosted in all regions except for Canada and Western Europe. Thailand and Malaysia are found to be the biggest indirect beneficiaries. They also present the overall increase (or otherwise) in household income across the world resulting from the deal, finding that Japan's welfare is reduced but Singapore's is increased. This they find is due to a "costly trade diversion, with increased imports coming from Singapore at the expense of lower cost suppliers elsewhere" (p. 21).

Francois, Meijl, and Tongeren (2003) have a far more opaque way of dealing with trade facilitation, presenting a range of estimates of the potential for cost reduction from studies including the European Commission, UNCTAD and APEC, before plumping for a modelled cost reduction of 1.5%. No explanation is given for how this number is arrived at but, as with Hertel et al., their CGE framework allows them to make interesting global estimates of the effects of a cost reduction, finding that LDCs in the Asia-Pacific region have most to gain from trade facilitation.

Minor and Tsigas (2008) also look at the effect of time delays in getting goods across borders on the ability of LDCs to trade, this time using a new database of time in trade costs per HS4-level commodity category. They are then able to forecast the impact of trade facilitation at a more granular level, looking at the effect of reducing time to export for sub-Saharan African targeting *its* particular mix of export goods. They find that the region would be able to diversify its exports, away from natural resources and towards

[6] The estimates of the cost of time at the border were first published as Hummels and Schaur (2012) but were available in unpublished form since 2000.

'light manufactures' such as footwear and furniture. This interesting and original study concludes on an optimistic note for the future of CGE modelling, including a side-swipe at CGE's great rival, the gravity model:

> The database created for this paper reinvigorates the role of CGE analysis in the area of estimating trade impacts. To date, analysis of trade facilitation's impacts on trade has been dominated by econometric estimates from gravity models, which rely on relatively simple data sets and correlations which provide limited insights... CGE models...are well suited to capture the full flavor of causal linkages. (Minor and Tsigas 2008, p. 19)

10.2.3 The constituents of export effectiveness

In this section we give a very brief overview of the categories of intervention that can lead to the kind of increase in export effectiveness which this chapter focuses on.

Bloom et al. (1998) discuss in the context of African countries the disadvantages that come from having large isolated inland populations, contrasting this with a high coastal population density. This they attribute to the fact that "isolated populations are not able to enjoy a high division of labor because many of the inputs needed for modern production must come from international trade" (p. 239). They also highlight the barriers effective infrastructure provision of low population densities and low rates of urbanisation, mentioning the stark fact that all of sub-Saharan Africa without South Africa had, in 1992, fewer kilometres of paved roads than did Poland, which had just 1.4% of the land area.

The finding that exporters whose population live primarily within 100km of a coastline grow faster than others is further supported by regression analyses by Radelet and Sachs (1998) and Gallup, Sachs, and Mellinger (1999). The former study also uses CIF/FOB ratios to measure transport costs and finds that they are significantly related to purchasing power parity-adjusted GDP, suggesting that poorer countries suffer higher transport costs for their exports. The latter study highlights the particular disadvantages faced by landlocked countries, suggesting that their coastal neighbours "harass land-locked countries, neglect the road networks that would link them to the coast, or impose punitive effective taxation through transit and port charges" (p. 213).

A similar CIF/FOB transport cost analysis is undertaken by Limão and Venables (2001) which, although not specific to LDCs is very widely cited throughout the LDC export effectiveness literature. They develop an innovative measure of infrastructure quality, calculating an average of a density measure of the road network, the paved road network and the rail network, along with the number of telephone main lines per person. This measure is found to correlate with both trade costs and trade volumes.

Clark, Dollar, and Micco (2004) also look at infrastructure, focusing on port efficiency and shipping costs, with reference to Latin American exports to the US. Beyond the obvious distance, and access to coastline, they look into the causes of the variable shipping costs seen across the region. They propose differing insurance costs, directional imbalances leading to empty containers for one leg of the return journey, absence of monopoly power and the quality of onshore infrastructure. Using various data from the World Bank, as well as from the World Economic Forum's Global Competitiveness Report, they regress a variety of these covariates against handling charges as a fraction of wages, and find that port efficiency is a strong determinant. This then leads them into an investigation of the correlates of having efficient ports, where they find that infrastructure quality and a perceived absence of organised crime are the strongest predictors.

In a study for the World Bank, Hausman, Lee, and Subramanian (2005) build on work by Subramanian and Arnold (2001) which finds that "differences in logistics performance are driven only in part by poor quality of physical infrastructure services such as road, rail, waterways, port services, and interfaces." They use survey data of freight forwarders, the middlemen who are often used by private firms in shipping products into and out of the country, and find that "institutional issues" such as the effectiveness of customs inspections and document processing are an even more important determinant of cost and time of shipments than the physical condition of road and rail infrastructure. They also construct an indicator of what they call "logistics friction" involving various measures of customs speed, border and port costs, bureaucratic complexity and availability of electronic (rather than physical) inspection facilities. This new indicator is then added to a gravity model which they use to estimate trade gains which would be available if such frictions could be reduced.

Eifert, Gelb, and Ramachandran (2008) show that regular power outages in national electricity grids in African countries leads to many African firms producing their own electricity at what they say is up to ten times the price of electricity from the public grid. They also use data from the International Telecommunications Union to show that broadband internet prices are "substantially higher than world averages" and cite work which suggests that "African firms pay higher bribes (as a percentage of sales) and lost a greater fraction of the value of their sales to crime and theft" (p. 1538).

Finally, Fu, Oum, and Zhang (2010) note that "like improvements of other shipping modes, the efficiency and quality improvements of air transportation promote trade and economic growth" and add that air cargo accounted for 40% of international trade by volume. Their study looks at the effects of air transport liberalisation, finding that is has led to "substantial economic and traffic growth" due to increased competition, efficiency gains due to route optimisation, and positive externalities to the overall economy due to employment opportunities, trade and tourism.

10.3 Modelling trade facilitation

In this section, we employ the exporter-specific intercepts from a regression model as a proxy for a country's overall export effectiveness, that is, all the unobserved factors which make a country a "good" or a "bad" exporter. As we have seen, there are many such factors, including export infrastructure, production capacity, trade tariffs, corruption, existing international relationships and currency strength.

By doing this, we will attempt to say something about the possible impact a trade facilitation policy such as "Aid for Trade" might have not only on the recipient, but on all countries in our model.

10.3.1 Estimating export effectiveness

Rather than seeking explicitly to model the factors which make a country a "good" or a "bad" exporter, we combine these effects into a unitless measure which we will refer to as "export effectiveness". This is a similar approach to the five-year averaged fixed effects employed by Helble, Mann, and Wilson (2012), reviewed above. Although the actual values of this measure are arbitrary, it will nevertheless be useful as a comparison between countries or, in future work, between time periods, or as a way to compare different scenarios.

In chapter 4 we estimated a simple gravity model attempting to explain trade flow magnitudes in order to estimate services trade flows. Here we will use the results of that analysis in a different context. Specifically, we will use specification (5) from table 4.1 and use the point estimates of the country dummies directly as our measure of "export effectiveness".

To recap the method of that earlier chapter, the gravity equation to estimate the flow of sector s from country i to country j is:

$$\log y_{ijs} = \beta_0 + \sum_k (E_k \delta_k) + \beta_1 d_{ij} + \beta_2 \log x_j + \beta_3 \log f_{is} + \beta_4 \log v_{is} + \varepsilon_{ijs} \qquad (10.1)$$

where $\delta_k = 1$ when $i = k$ and zero otherwise, E_i is the export effectiveness measure we are estimating, d_{ij} is the shortest distance between i and j (i.e. neighbouring countries have $d_{ij} = 0$), x_j is the total production of the importing country, f_{is} is the final demand for sector s in the exporting country, and v_{is} is the per-unit value added. The dependent variable, y_{ijs} is a set of commodity trade flows from UN COMTRADE for a particular year.[7]

[7] In this particular case we repeat the analysis of chapter 4 but this time with data from 2010.

Table 10.1: A summary of the "export effectiveness" (EE) measure describing how attractive a country is an an export partner beyond the effects included in the regression specified in section 10.3.2. The measure is unitless and runs from 6.5 for the USA to -0.6 for Bhutan.

(a) Most effective exporters

Country	EE
USA	6.5
China	6.2
Malaysia	6.1
Germany	5.9
Singapore	5.6
France	5.6
UK	5.5
Italy	5.4
Japan	5.3
Korea	5.2

(b) Least effective exporters

Country	EE
Bhutan	-0.6
Eritrea	-0.5
Iraq	-0.5
Cape Verde	-0.3
Nepal	-0.3
Haiti	-0.2
Angola	-0.1
Chad	-0.1
Maldives	-0.0
Mauritania	-0.0

Table 10.1 shows a summary of the unitless measure of export effectiveness as estimated by equation (10.1). The measure ranges from -0.6 to 6.5, with the worlds least-attractive trading partner being Bhutan, and the most attractive being the USA. It is important to note that this is not merely a size measure: the inclusion of f_{is} absorbs the fact that the USA is the largest economy in the world.[8] China and Malaysia follow in the list of top exporters in 10.1a. This could provisionally be linked to what Stiglitz (2007) refers to as "the Malaysian miracle," of successful poverty reduction, job creation and social cohesion. Germany is Europe's most attractive exporter, followed by France, the UK and Italy. Singapore, Japan and South Korea also score highly in export effectiveness terms.

Of the top 10 least attractive export partners in 10.1b, Bhutan is together with Nepal in being hugely dependent on a single market (India in both cases) for its exports. Other very low export effectiveness countries are war-torn (Eritrea, Iraq, Angola), land-locked (Chad) or island nations (Cape Verde, Haiti, Maldives).

10.3.2 Modelling Approach

Our model assumes that export facilitation policies have some kind of positive effect on the export effectiveness measure estimated in section 10.3.1. It is conceivable that we might estimate the effect of a dollar of Aid for Trade on export effectiveness by

8 Note also that total production of the exporter cannot be included in the regression since it would be co-linear with the country indicator δ_k.

using time series of Aid for Trade payments and an equivalent series of estimates for export effectiveness, but the challenges of identifying Aid for Trade from total official development assistance (ODA) seem formidable and it might be hard to justify an assumption that *all* aid attempts to affect the ability of the recipient to export.

Instead, we will restrict ourselves to a comparison of the effect of an arbitrary, but fixed, increase in the export effectiveness of each aid recipient country. This will still allow us to say something about the elasticity of GDP with respect to export effectiveness, and hence the countries which have the potential to most benefit from increased export effectiveness (and, equivalently, which countries stand to lose the most from such increases elsewhere.)

Our analysis will focus on a subset of aid recipients. Namely, the recipients of "ODA and official development aid" as a fraction of GDP greater than 2% in 2010, according to the World Bank's World Development Indicators, including only "large countries", by which we mean those with a population of over 2 million.[9] This restricts us to 51 countries of which 31 are in Africa.[10]

For each of these 51 countries we will increase the export effectiveness by a small fixed amount, ε, and solve equation (10.1) with the error term set to zero and the various β parameters as estimated by the regression. In other words, we calculate the estimated trade flows, $\hat{y}_{ijs}$, for each exporter, importer and sector first with E_k and then with $E_k + \varepsilon$. Then, using these $\hat{y}_{ijs}$, we will calculate import propensities using the same formula used to calculate them in chapter 3, which simply divides each flow by the total imported by a given importer. These import propensities are then inserted into the demonstration model and the model is recalculated. This allows us to track the change in GDP in every country in the model for a change of ε to the export effectiveness of each of our 51 recipient countries in turn.

As outlined in the setup of the demonstration model, countries with no imports of a particular sector in the COMTRADE data upon which the model is based cannot calculate their import propensities in this way, and so it will be assumed that they import all their domestic requirements for that sector from the Rest of the World (RoW) entity.

9 Many of the largest recipients of ODA *per capita* are small Pacific islands and other countries with smaller populations.

10 The non-African "large" aid-recipient countries are: Afghanistan, Albania, Armenia, Bolivia, Bosnia Herzegovina, Cambodia, Georgia, Haiti, Honduras, Jordan, Kyrgyzstan, Laos, Mongolia, Nepal, Nicaragua, Papua New Guinea, Moldova, Tajikistan, Vietnam and Yemen.

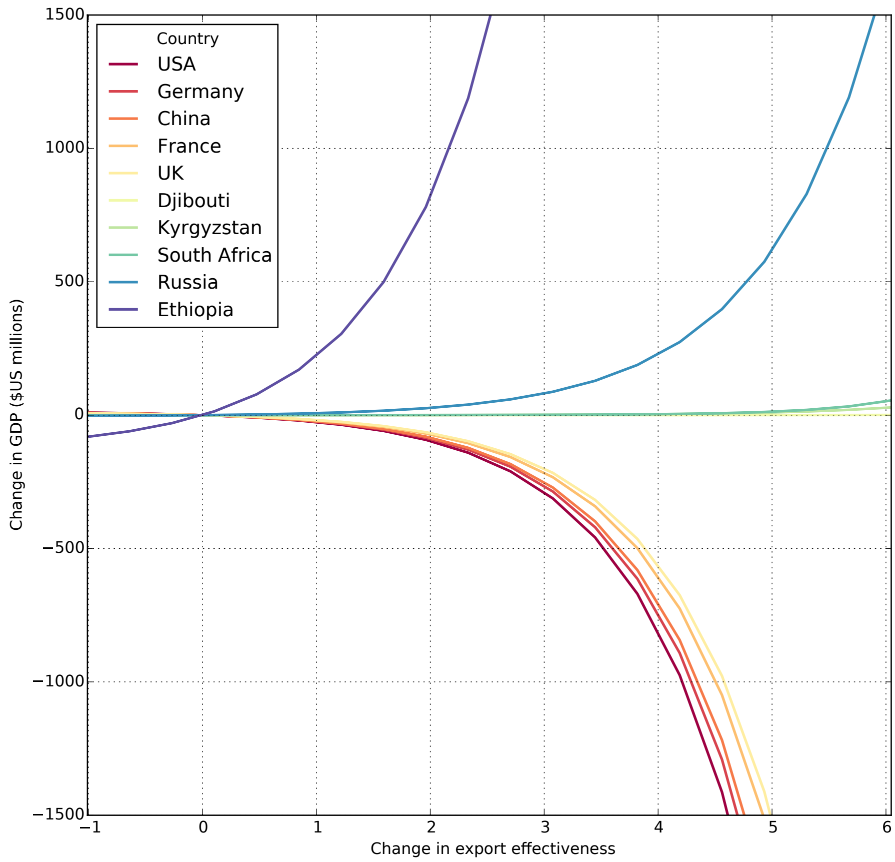


Figure 10.1: Changes in GDP for the five most positively affected and five most negatively affected countries due to a change in export effectiveness of Ethiopia.

10.4 Results

Throughout this section, we will consider the results of the experiment described in section 10.3.1, using $\varepsilon = 0.1$.

10.4.1 Non-linearity in ε

As the choice of $\varepsilon = 0.1$ was arbitrary, it is important to understand how the model response depends on this parameter choice. By picking a particular recipient and calculating the effects on the model at various values of ε, we can understand how the parameter choice affects the outcome.

In figure 10.1, we pick Ethiopia (export effectiveness 0.42) as our recipient country and vary ε such that Ethiopia's export effectiveness varies across the whole scale of modelled countries, from -0.6 of Bhutan to +6.5 of the USA. The changes to GDP of the five most increased and five most decreased countries are tracked as Ethiopia's export effectiveness increases from that of Bhutan to that of the USA.

Table 10.2: The effect of increasing export effectiveness by 0.1 on the affected country itself. Effect is measured in terms of change in GDP, both absolute and percentage. The top ten by percentage are shown.

Country	GDP ($ US millions)	Change	Change (%)
Vietnam	40,890	2,163	5.59
Togo	1,856	33	1.82
Honduras	6,209	92	1.50
Republic of Moldova	2,406	35	1.46
Papua New Guinea	2,668	37	1.40
Ivory Coast	10,652	123	1.17
Nicaragua	5,191	54	1.04
Senegal	7,871	73	0.93
Bolivia	7,183	65	0.91
Cambodia	2,754	23	0.86

As expected, Ethiopia itself is the country most positively effected, followed by Russia, South Africa, Kyrgyzstan and Djibouti. Negatively affected by Ethiopia's improved export effectiveness are Germany, the Netherlands, the UK, France and Belgium. Clearly the effects on each of these countries is highly non-linear but, crucially for our purposes, appear to be monotonic, with an increasingly steep slope. This is relevant because it means the lines in figure 10.1 do not cross as changes in export effectiveness become more extreme. We can therefore say that conclusions drawn about the relative, ordinal effects of export effectiveness at a particular level of ε will hold for any level of ε.

10.4.2 Own-country effects

We first consider the approach taken by the typical study of the effects of aid: the effect on the recipient country itself.[11]

Table 10.2 shows the results of the experiment outlined in section 10.3.2 on the country whose export effectiveness was increased. Listed are the ten most improved countries by percentage change to GDP. By far the most improved country is Vietnam, increasing its GDP by 5.6% in response to the increased export effectiveness. The remainder of the ten most improved countries are more closely spread, with just a single percentage point between Togo, at number two, and Cambodia at number ten.

It is interesting to note that while three-fifths of the 51 countries in our recipient sample are African, only three of the top 10, Ivory Coast, Senegal and Togo, are in Africa.

[11] Recall though that a global economic model is being solved for each change in export effectiveness. Thus these results include input-output multiplier effects and any effects from changed levels of trade throughout the model.

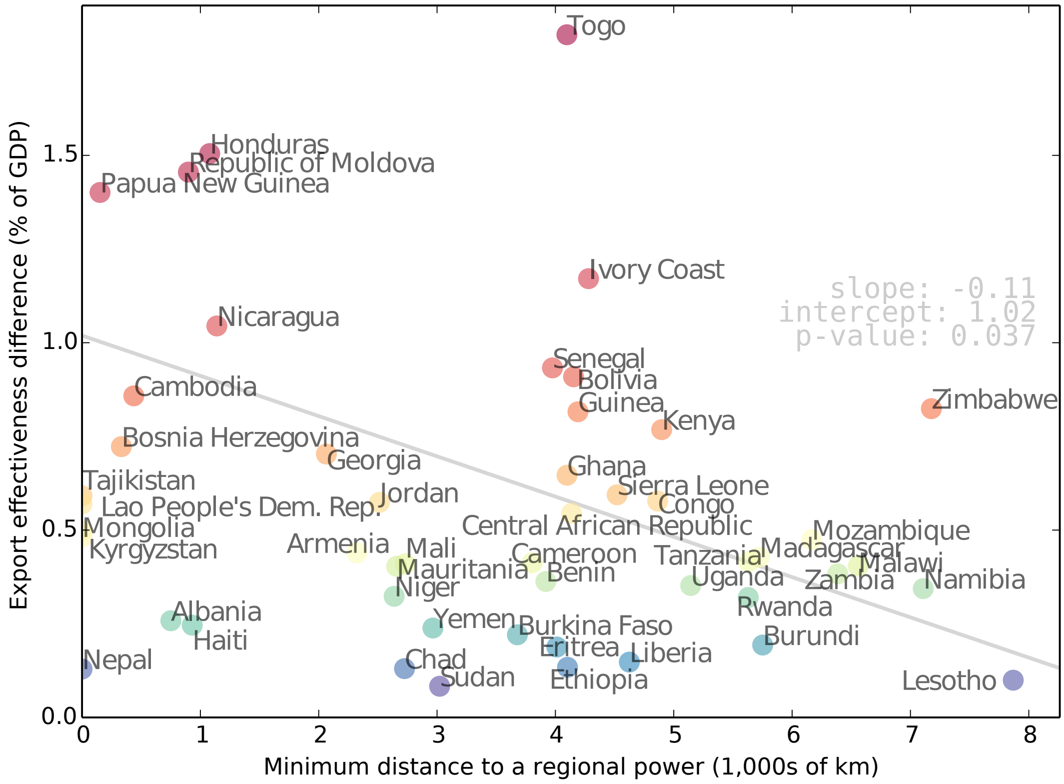


Figure 10.2: Scatter plot showing the relationship between the minimum distance of a recipient country to a regional power (USA, China, Germany or Australia) and the results of the export effectiveness experiment outlined in section 10.3.2. Vietnam is a huge outlier, with an export effectiveness difference of 5.6 and minimum distance (to China) of zero. It is not shown on the graph, but did contribute to the regression line. See p. 276 for a discussion of the Vietnam case.

The others are south-east Asian: Vietnam, Cambodia; or Latin American: Honduras, Nicaragua, Bolivia.[12]

This raises the question of whether a recipient country must be geographically near to a regional power in order to benefit from an export effectiveness increase. To test this, we regress the percentage change from table 10.2 against the distance from either the USA, Germany, China or Australia, whichever is the closest. Figure 10.2 shows a scatter plot of the minimum distance, in thousands of km, from one of these regional powers, against the own-country benefit from an export effectiveness increase.

The relationship has the expected negative sign (-0.11) and is significant at the five percent level ($p = 0.037$). Visually, the relationship is rather weak and there are some very noticeable outliers.

[12] The top 10 also includes Moldova and Papua New Guinea.

Among those countries whose export effectiveness results are less positive that we might expect given their proximity to a regional power, the largest outlier is Nepal. This country is land-locked between India and China and, as Jayaraman and Shrestha (1976) point out,

> [Nepal's] excessive dependence on one market for its entire foreign trade places Nepal in a unique problematic situation. While the small size of the country impels Nepal to resort to international trade, its geographical position limits its option to trade with countries other than India.

Although improved since the 1970s, this situation is still severe, with India making up over 50% of Nepal's exports in 2012 (Simoes and Hidalgo 2011). This reliance on a single export partner shows why Nepal is unable to benefit from increased export effectiveness: no other market is available, so being a more attractive export partner is of little use. This same effect may explain Haiti's presence as a negative outlier (85% of Haitian exports went to the USA in 2005), Chad (75% to the USA) and Sudan (83% to China and Japan combined).

On the positive side, the country to most benefit despite its comparative distance to a regional power is Togo. This country is such a stark outlier that it may deserve some closer inspection. Table 10.3 shows a summary of the results for Togo. The biggest increase is to the USA with $1,280 additional trade. This trade with the USA is broken down into sector in table 10.3b. The sector to benefit most is Mining, with $202 additional export (Togo exported $42,000 of phosphates to the USA in 2005), followed by Electricals (around $130,000 of integrated circuits was exported to the USA in 2005) and Vehicles ($1,500 of HS code 870590, "special purpose motor vehicles", which seem primarily to be road manufacturing vehicles, such as bitumen spreaders.)

Togo is not a particularly large exporter to the USA, with only $11 million in 2005, making it the 35th largest of 53 African countries. Togo's large gains seem also not to be caused by inappropriate estimates for the value added per unit of Togo's sectors: the sector with the highest value added per unit is construction with 0.35 which has the same order of magnitude as Ethiopia's construction sector, with 0.32, and that of the USA which has 0.48. Ethiopia fares much worse in the export effectiveness experiment than does Togo, which is why this a useful comparison. But unlike the negative outliers, Togo's export profile was surprisingly diverse in 2005. The largest trade partner, Ghana, had only 15% of Togo's total exports.

Looking at the Togo result, we might hypothesise that benefiting from export effectiveness increases requires a broad spread of export partners. To test this, we calculate a Herfindahl index of export-partner concentration based on the share, s_{ij}, which country

Table 10.3: Increases in trade flows from Togo, due to the export effectiveness experiment, in which Togo's export effectiveness was increased by 0.1.

(a) per importing country

Importer	Trade ($US)
USA	1,280
France	716
Germany	598
UK	447
Italy	445

(b) per sector to the USA

Sector	Trade ($US)
Mining	202
Electricals	185
Vehicles	179
Chemicals	153
Textiles	122

j takes of country i's total exports. Throughout what follows we will do all calculations based on 2005 data.[13] The concentration measure is

$$H_{ij} = \sum_{j=1}^{C} s_{ij}^2 \tag{10.2}$$

where C is the number of countries in the data (in our case, there are 230 trade areas with an ISO3 country code, and we will use this as C throughout the following.) The measure ranges from $1/C$ for a country with perfectly even export partners, to unity for a country with only one export partner.

We can normalise H to run between zero and one by a simple adjustment:

$$H_{ij}^* = \frac{H_{ij} - 1/C}{1 - 1/C} \tag{10.3}$$

Figure 10.3 shows a scatter plot of the relationship between H_{ij}^* and the results of the export effectiveness experiment. The relationship has the expected negative sign, indicating that countries with more concentrated export markets did worse on average, but the relationship is not significant. This may be because many countries have very similar values of H_{ij}^*, with the majority being in a cluster around 0.1 to 0.3. Such relationship as there is seems to be caused mainly by outliers such as Vietnam, Togo, Moldova, Haiti, Lesotho and Chad.

10.4.3 Positive other-country effects

Although the majority of effects were negative for countries other than those having their export effectiveness increased (hereafter "other" countries), a few other countries

[13] Note that here a negative flow from Gabon to the USA has been corrected to a positive flow. Checks with online resources have been made to ensure the feasibility of this. This was done as a special case since Gabon has few enough trade records that one error is likely to seriously impact the results.

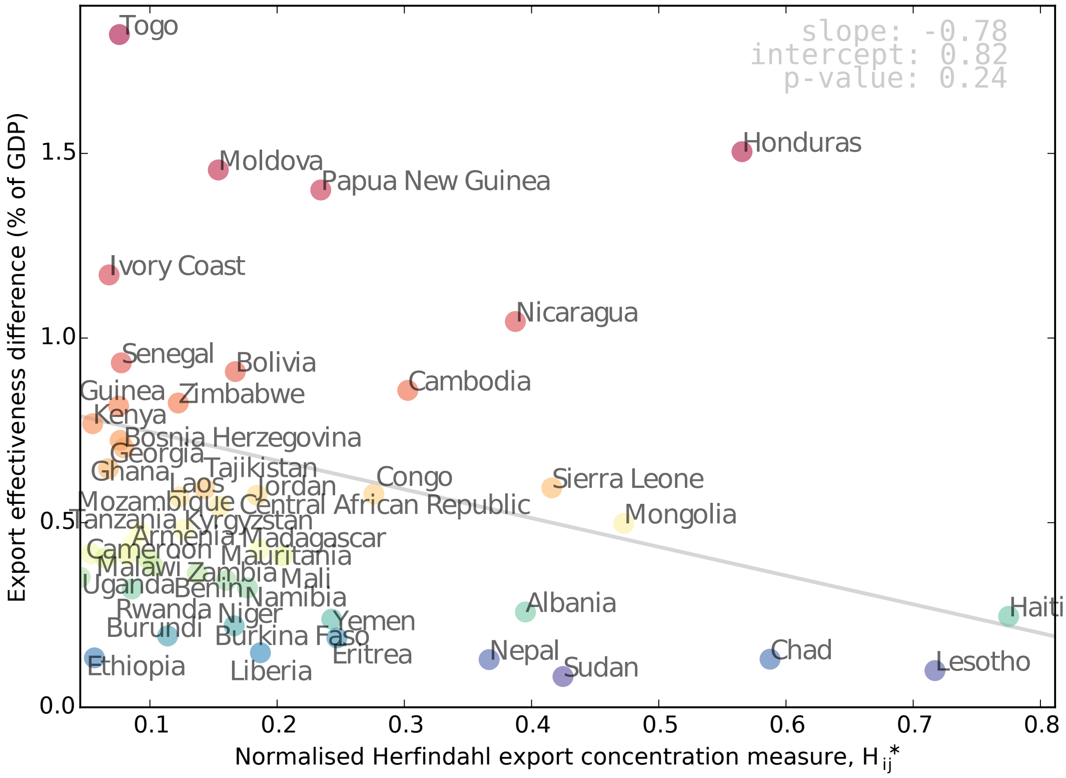


Figure 10.3: Scatter plot showing the relationship between the normalised Herfindahl index of export-partner concentration and the results of the export effectiveness experiment outlined in section 10.3.2. Vietnam with an export effectiveness difference of 5.6 and an H^* of 0.07 is not shown on the graph, but does contribute to the regression line. See p. 276 for details of the Vietnam case.

benefited. Table 10.4 shows a summary of the positive results for other countries. The export effectiveness of the country in the “Recipient” column was increased as per the experiment of section 10.3 and positive GDP effects on other countries were measured. The sum of these GDP increases is reported in the “Total effect” column, and the ten largest by total effect are shown. The following two columns show the largest three other-country benefactors of the increased export effectiveness in the Recipient country.

The largest positive effect by a large margin is that associated with Vietnam, with a total effect of $118 million. The entire positive effect is due to an increase in GDP in Korea. This is an interesting result because Korea is only the fourth largest importer to Vietnam (11%), after China (18%), Singapore (13%) and Japan (12%) and because Vietnam makes up only 1.2% of Korea’s total export market. Figure 10.4 shows a part of a network representation of all the changes resulting from an increase in Vietnam’s export effectiveness. The nodes of the network are either countries, labelled in light blue, or country/sectors (CSs), labelled in black as “sector (country)”. All edges represent a change: those between CSs show changes in trade flows, and those between countries and CSs represent changes in total production of that sector in that country. Nodes are sized by total degree (in- plus out-degree). The width of the edges is proportional to the logarithm of the magnitude of the change, with positive changes shown in blue and negative changes shown in red. For visual clarity, edges with weight smaller than 3.25 have been filtered out. Additionally, only those nodes which are neighbours of the “VNM” country node to a depth of 3 or less are shown. The sectors shown are the seven “super-sectors” described in table 5.2 in chapter 5.[14]

We can see immediately that the change to total production in Vietnam (VNM, slightly left of centre of the image) is large and positive, as we would expect. There are large increases in the “raw” and “primary” sectors, which then trade with China, Japan and the USA. This leads to modest reductions in the production of these sectors in these countries. In Korea, in the top-left of the image, the “primary” and “secondary” sectors are reduced in output, but the “trade”, “services” and “transport” sectors all see comparatively large increases. These sectors all have increased trade with Vietnam. Russia and Canada also see increased output in the “transport” sector due to increased trade with Korea. The “raw” sectors in Thailand (THA, left-centre of the image) and Singapore (SGP, centre) also see increased production due to increased trade with Vietnam.

This diagram is presented as an indication of the complexity of dealing with global models of this kind. As a great number of the sectors are inter-related internationally, many thousands of outputs change in response to each input change and it becomes difficult to trace effects back though the model. A diagram such as this, while

[14] See the footnote in section 5.2.2 for a caveat about the nature of this categorisation.

Table 10.4: Other-country positive effects from the export effectiveness experiment. The country in the "Recipient" column had its export effectiveness increased by 0.1. The "Total effect" is the sum of all the positive changes in GDP from countries other than the recipient itself. The ten largest by this measure are shown. The top three other-country effects are shown per recipient country.

Recipient	Total +ve effect ($US millions)	Other country	Δ GDP ($US millions)
Vietnam	118.25	Korea	118.25
Honduras	18.93	Canada	13.52
		USA	4.98
		Israel	0.42
Kyrgyzstan	11.95	China	8.37
		India	1.75
		Turkey	0.34
Albania	9.65	Russia	9.62
		Kyrgyzstan	0.03
Kenya	9.01	Russia	7.96
		Kyrgyzstan	0.58
		South Africa	0.43
Papua New Guinea	8.97	Australia	7.93
		Indonesia	1.04
Georgia	8.64	Russia	8.54
		Azerbaijan	0.08
		Kyrgyzstan	0.01
Congo	8.32	Russia	3.60
		Israel	2.54
		South Africa	1.83
Jordan	6.68	Russia	5.71
		Kyrgyzstan	0.89
		Saudi Arabia	0.08
Bolivia	6.40	Chile	3.99
		Russia	2.41

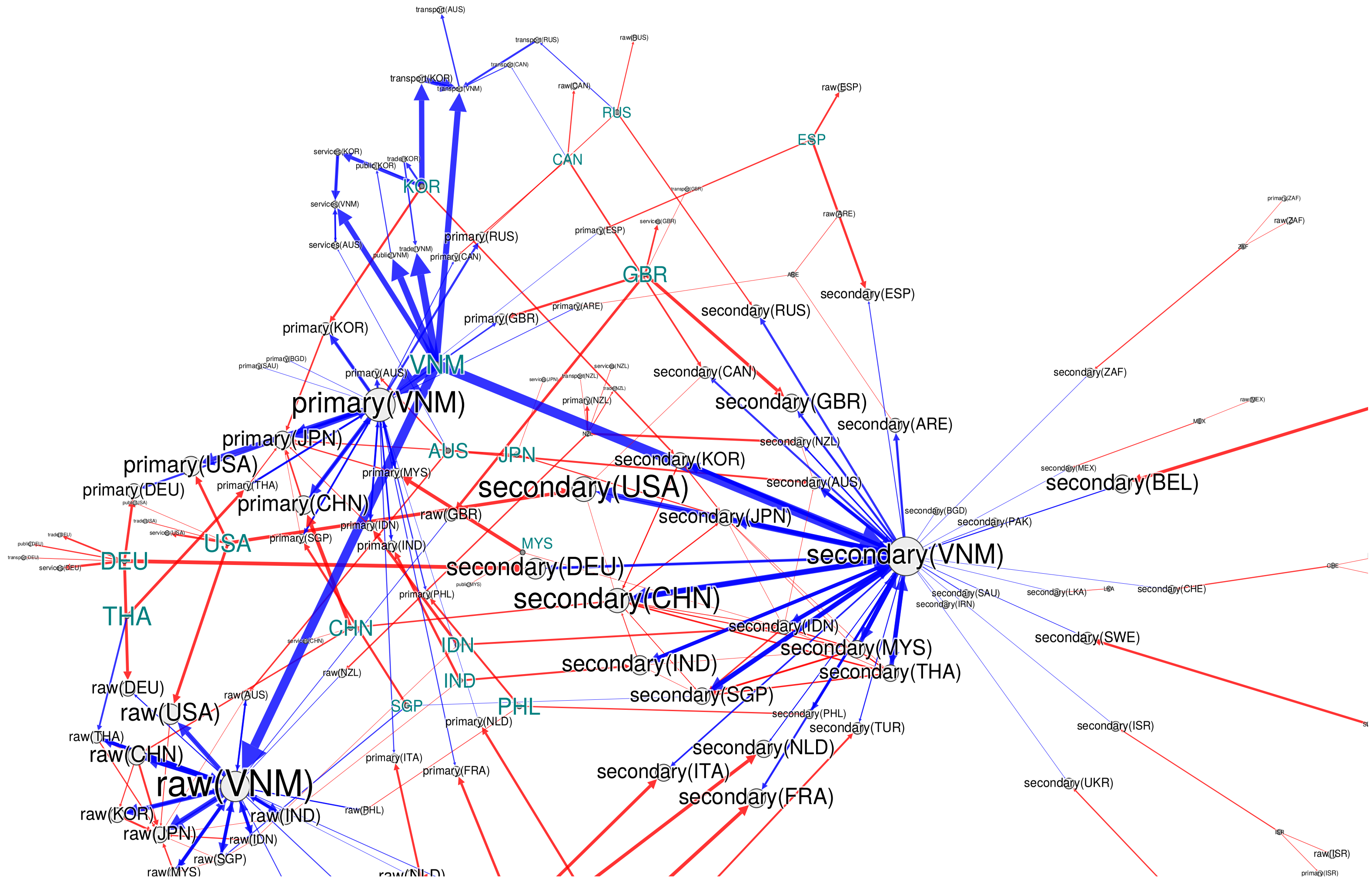


Figure 10.4: A part of a network of changes resulting from an increase in export effectiveness in Vietnam. Blue arrows show increases and red arrows show decreases. Nodes are either countries (labelled in light blue) or country/sectors (CSs, labelled in black). Edges between CSs represent changes in trade flows. Those between countries and CSs represent changes in total production. All changes are logarithmic. Sectors are aggregated according to the classification on p. (144)

still complicated-looking, can guide us in beginning to trace key routes through which mechanisms are having their effects.

As with all of the analysis in this thesis, we present this particular Vietnam case study as a methodological and visualisation example. It is important to see the additional insight that can be brought by analysing problems with the surrounding global economic context. And that the additional complexity this brings requires innovative visualisation tools to help us make sense of the results.

10.5 Choosing an aid recipient

There is a risk in studying the effect on the global network of changes in export effectiveness, that we are merely focusing on countries because of the overall size of their economy, not for some fundamental property of their production technology and position within the trade network. Perhaps a better metric, and one which may allow us to assess the suitability of a particular recipient country for receipt of trade facilitation assistance, would be one which explicitly compares own-country gains to other-country losses.

10.5.1 A metric for assessing bilateral aid decisions

Here we focus on the perspective of a particular donor country, seeking to optimise a trade-off between the improvement to the GDP of a recipient against the reduction in GDP of the donor country itself. We assume throughout that the actual aid itself is insignificant to the GDP of the donor. Although some donor countries have committed to spending 0.7% of GDP on foreign aid, it is safe to assume that any one aid flow, particularly that part of it dedicated to trade facilitation, is much smaller than this as a percentage of GDP.

We will define as a donor the top 15 largest overall givers of aid in 2010 which were, in descending order, the US, the UK, Germany, France, Japan, the Netherlands, Spain, Canada, Sweden, Norway, Australia, Saudi Arabia, Belgium, Italy and Denmark.[15]

The idea that aid is not always allocated based on altruism is discussed in the literature (Dudley and Montmarquette 1976; Trumbull and Wall 1994), with donors having political motivations (McKinlay and Little 1977; McGillivray and Oczkowski 1992), or expecting additional trade benefits, including the finding that “OECD countries allocate more aid to recipient nations who import goods in which donor nations have a comparative advantage in production” (Younas 2008, p. 661). The United Nations Development Programme described the allocation of aid as being ‘strange and arbitrary’ going as far as to say “[development aid] is determined not by the needs of developing countries, but by the fluctuating goodwill of the people and their parliaments in the rich countries. As a result, it is largely ad hoc and unpredictable” (UNDP 1992, p. 45).

We will therefore assume that each donor has a mix of altruistic and selfish motivations. It cares about recipient gains in a linear fashion, but cares about its own gains in an exponential fashion. This donor preference will be denoted p_{ij}, meaning the preference

[15] These are the largest givers of aid according to the data series “Total flows by donor (ODA+OOF+Private)” from the OECD website, retrieved 4th September 2015.

of donor i for giving to recipient j. Gains/losses in donor i due to an export effectiveness increase in recipient j are denoted δ_{ij} and changes in the recipient country, which are not dependent on the donor, as δ_j. We define our preference function as:

$$p_{ij} \equiv \delta_j e^{\delta_{ij}} \tag{10.4}$$

This formulation has several attractive properties. Firstly, since recipient countries always benefit from an increase in export effectiveness, and since the exponential function is always positive, we have $p_{ij} > 0 \ \forall \{i, j\}$. Secondly, positive δ_{ij} has a (much) greater effect on p_{ij} than positive δ_j. And finally, negative δ_{ij} diminishes the impact of a positive δ_j. These last two make precise the notion that donor countries are selfish.

By ranking recipient countries per donor country according to their p_{ij} donor-preference score, we can see which countries make "good" recipients for each donor. Table 10.5 shows the rankings for all donors and shows those recipients which rank inside the top three for at least one donor. We see a mixed picture emerging.[16]

Kyrgyzstan is the most preferred recipient for a large number of donors. But a number of large donors (including the USA, the UK, and Canada) buck this trend. The top three for the US are Honduras, Nicaragua and Kyrgyzstan, with Honduras also ranking highly with Canada, Norway, Denmark and Saudi Arabia. The top three for the UK are Laos, Cambodia and Nepal, with Laos also being a good recipient for France, Italy and Germany. We thus begin to see the possibility that there are clusters of donor/recipient relationships which go beyond the simple answer that Kyrgyzstan is the most popular.

Figure 10.5 attempts to show some of the more subtle aspects of the relationship between donor preference and recipient country. The nodes in the figure are of two sizes: large for a donor and small for a recipient. The edge weights are proportional to the inverse of the within-donor rank of the donor preference as calculated by equation (10.4). Only those edges relating to ranks of 10 or less are shown. The three colours on the graph represent membership of three clusters, found by a modularity-maximisation clustering algorithm due to Blondel et al. (2008) with a resolution of 1.5 (chosen to result in three clusters).

We clearly see a cluster of the USA, Canada and Norway around Nicaragua and Honduras; one of Germany, the UK, Italy and France around Laos, Mongolia and Nepal; and a large cluster around Kyrgyzstan. The large cluster could potentially have a subcomponent of the Netherlands, Belgium, Sweden and Australia around Papua New

[16] Since these are ranks, clearly a lower number is "better." Thus a rank of 1 means the recipient is the donor's most preferred.

Table 10.5: The ranking of the donor-preference term p_{ij} for the 15 largest overall givers of aid in 2010, given by equation (10.4). Only recipient countries which appear in at least one donor's top 3 are featured.

Donor	Bolivia	Cambodia	Cameroon	Honduras	Ivory Coast	Kenya	Kyrgyzstan	Laos	Mongolia	Nepal	Nicaragua	PNG	Senegal	Sudan	Uganda
Australia	11	20	7	17	4	3	2	32	41	38	15	1	5	31	18
Belgium	4	3	16	40	45	43	1	10	32	29	6	2	20	24	7
Canada	42	18	8	1	3	6	45	27	37	33	4	17	2	28	11
Denmark	2	17	8	3	12	6	1	28	39	34	4	7	5	31	19
France	43	3	36	48	49	47	1	2	6	4	42	8	46	7	27
Germany	44	19	41	47	49	48	1	2	3	4	42	25	46	7	29
Italy	40	6	41	48	49	47	1	2	9	8	38	4	46	3	28
Japan	44	33	3	49	46	40	1	26	30	31	47	42	39	15	2
Netherlands	25	3	34	45	49	46	1	4	21	17	22	2	42	14	7
Norway	5	18	9	2	1	3	7	28	39	36	6	12	4	32	19
Saudi Arabia	6	19	9	3	2	1	10	31	39	38	8	14	4	32	18
Spain	8	6	12	43	24	39	1	18	33	31	9	3	2	26	13
Sweden	2	15	6	11	24	8	1	27	39	34	4	3	10	31	17
UK	43	2	40	48	49	47	25	1	5	3	42	6	46	8	26
USA	47	36	40	1	48	49	3	21	17	16	2	43	44	22	29

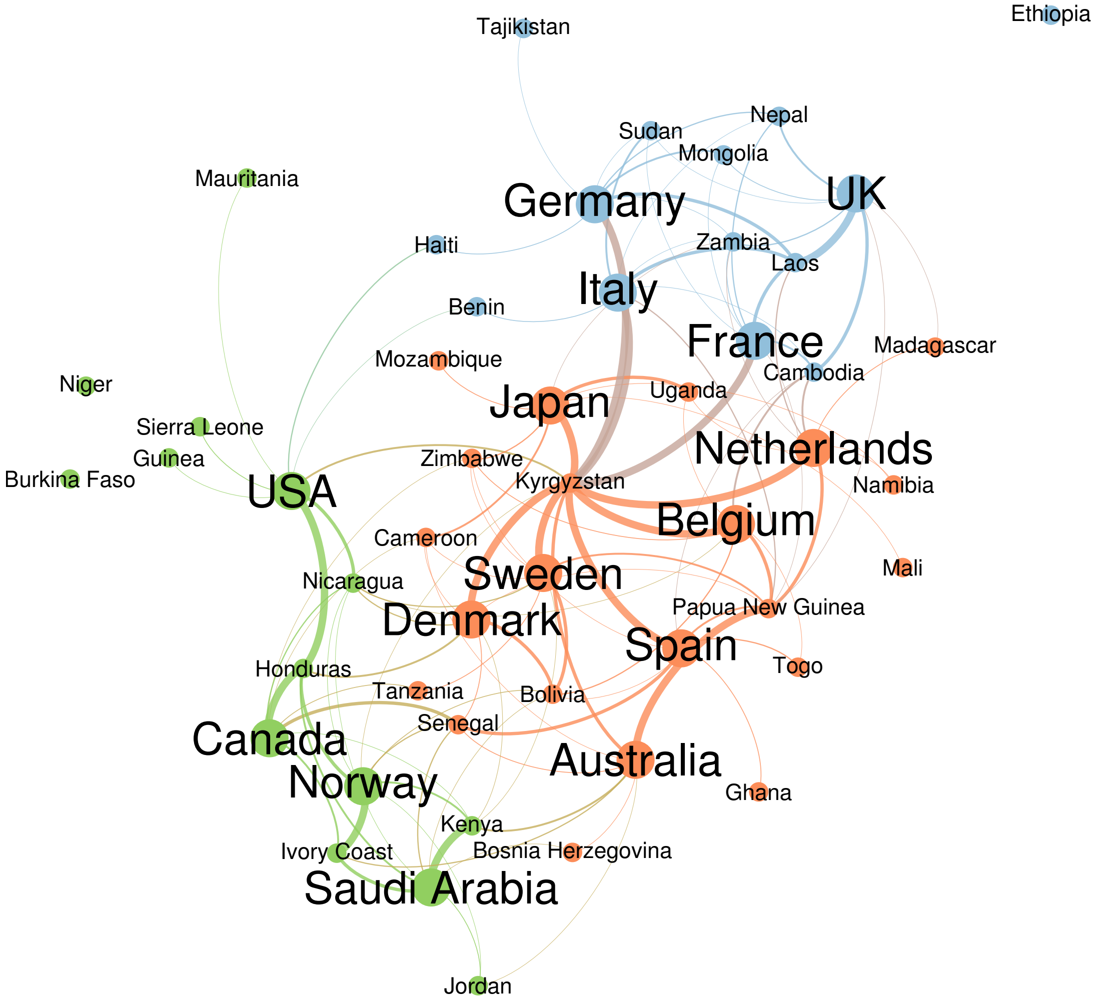


Figure 10.5: Donors and their preferred recipients according to a partly selfish utility function. Large nodes are donors and small nodes are recipients. Edge weight is proportional to the inverse of the rank of the donor preference for that recipient, as calculated by equation (10.4). Node colour represents membership of a cluster, based on a modularity maximisation algorithm.

Guinea. The results of the clustering add weight to our suspicions that donors congregate around different “favoured” recipients

10.6 Conclusion

In this chapter, the last of our three social science extensions to our model, we have included a wide range of analyses, all focused on the effects of improving export infrastructure in a sample of large, aid-recipient countries.

We used a gravity-style regression model (taken from work done in chapter 4) to estimate a set of exporter-specific intercepts which capture all factors which affect a country's ability to export beyond its geographical distance to trade partners, its total production, its final demand and a measure of its production efficiency. We labelled this catch-all measure "export effectiveness."

We then ran an experiment which increased export effectiveness by a small amount in each of our sample countries and measured the effect on GDP of every country in the model. By doing this, we could pose two categories of question. The first is of the type typically asked by these kinds of analysis: which country benefited the most from an increase in its export effectiveness?

But the second category of question is unique to a global modelling context: what other effects did an increase in export effectiveness have? Specifically, who might *lose* from such an increase? Since improved export efficiency *per se* generates no new final demand, increased flows to the country in question must be accompanied by reduced flows to other countries.

It is this last idea which led us to our final analysis of the chapter, and one which is more explicitly to do with aid. By measuring the positive effect of increased export effectiveness against the negative effect to each donor country, we were able to estimate which recipients each donor might favour, given a selfish preference function which values recipient gains and donor losses, but the losses more sensitively than the gains. By doing this, we were able to group similar donors in terms of their preferred recipients.

As with the other analyses presented in this thesis, the work done here is intended as a demonstration of what is possible when a model is global and data-driven, rather than heavily theoretical or country-specific. Even if the results here are preliminary at best, we can see the new kinds of discussion we can have by adding the international context of a global economic model.

Chapter 11

Thesis discussion and final conclusions

11.1 Introduction

The work within this thesis has introduced a simple model of the global economy which combines input-output models of national economies and an international trade model. Several demonstration analyses have been performed in several different contexts, with the model in its basic form and extended to include ideas, data and modelling approaches from other social sciences.

This chapter now concludes the thesis, first by giving a brief summary of the major findings in section 11.2. Then, in section 11.3, by critiquing the methodology used in the design of the model, as well as presenting some possible defences against those critiques. And finally, with some suggestions for the direction future research and applications might take, in section 11.4.

11.2 Summary of findings

We set out in the introduction to show how the addition of global context, and the use of innovative visualisation and analysis techniques, can shed new light on existing studies of human dynamic systems. The model was a half-way house between the rigidity of input-output and the formal technical complexity of a CGE/GTAP approach. In contrast to input-output modelling we separated the national and the international levels of the model, allowing the possibility in future of using a more standard, non-linear trade model. We introduced a simple gravity-type model in chapter 4 and employed it in chapter 10 where it allowed us to estimate how the global system responds to an increase in the export attractiveness of a given country. What was a simple exercise in adjusting import propensities here would be complex to achieve using traditional MRIO, and hard to insert into the rigorous theoretical framework of CGE.

With the introduction of import propensities and import ratios in chapter 3 we showed how countries could be added to the model with a greatly reduced requirement for data (see p.78) than in MRIO. This facilitated the addition of around 150 countries to the initial model based on the 40 countries of the world input-output database (WIOD), work carried out by Thomas Oléron Evans which we analysed in detail in chapter 5.

We then embarked on a series of analyses. The first of these, in chapters 6 and 7, involved analysis of the model itself, without recourse to additional data. These we called "within-model" analyses. In chapter 6 we looked at the sensitivity of the model to demand shocks, finding that global output was most sensitive to a shock in the Chinese vehicles sector. This led us to a discussion of economic self-sufficiency, which we found was closely related to country size, with a few notable outliers.

We then turned our attention to import substitution, showing how import ratios could be used as a simple way to model changes in a country's import dependence. Here, we found that substitution in different sectors had very different foreign and domestic effects, finding that the US mining sector had the biggest domestic benefit from import substitution as a ratio of losses caused to the rest of the world.

Our second group of within-model analyses, in chapter 7, focused on changes to patterns of final demand. We developed a mathematical method of adjusting one country's final demand (the "source") in such a way that it approached that of another country (the "target"). Here we also introduced a way of thinking about the model as a network, including novel use of some of the tools of graph theory to analyse the relationship between sources and targets. By recording which targets were beneficial to which sources in terms of increases to GDP, we were able to construct a matrix of improvement relationships which we viewed as the adjacency matrix of a "network of improvers." We then took this network of improvers and ran some typical graph-theoretic analyses on it, finding that improver relationships formed into two clusters, one of rich countries, and one of poorer countries, suggesting that these groups of countries might have different targets to which they should aspire, in order to maximise their GDP.

We then came to our suite of three social-science analyses which expanded the focus beyond the basic model to include data from other areas of study. These areas of study were migration, in chapter 8, international security, in chapter 9 and international development in chapter 10.

In chapter 8, estimates of global migrant flows from the work of Adam Dennett were used as a starting point for a number of experiments on the economic effects of migration. We showed how the commonly-held view that migration increases bilateral trade created both benefits and costs across the global system. This analysis concluded that the USA was the biggest winner from global migration patterns followed by Australia, and that Mexico suffered the greatest cost, along with Portugal, Romania and Bulgaria. As will all our results, we stressed the novelty of the modelling and visualisation approaches, rather than the real-world implications of these findings.

Chapter 9 on international security took as its starting point a set of threat "flows" from the work of Peter Baudains, based on alliance similarity and the level of military spending. With some very simplified assumptions about the economic impact of military spending, we showed how one might add global context to security studies, and how to model the economic gains made due to the network of security threats present around the world. Our unexpected finding was that the West-African nation of Guinea had the highest percentage benefit to GDP, due to its important role in exporting metals to Russia, although we also noted that Baudains' threat measure seemed to overstate the importance of military spending, leading to some unusual results.

Our final social science extension was in the context of economic development and sought to discover the wider implications of programmes to improve export effectiveness. We saw that it was possible for third-party countries to lose out when one country became a more effective export partner.

We also showed a method for quantifying which recipient country a selfish donor would most prefer to assist in this way. To do this, we made the assumption that aid from a donor country had the effect of improving the export effectiveness of the recipient country, without generating any additional demand worldwide. This then allowed us to see which countries stood to lose and which to gain from an increase in export attractiveness in each recipient country. By postulating a simple self-regarding preference function, we were then able to rank recipient countries in terms of their own ability to benefit from aid, set against the losses caused to the donor from any export demand the newly attractive recipient might usurp.

Throughout this summary, we have seen that our global, data-driven philosophy has allowed us to present concrete, real-world answers to the research questions we have posed. Our purpose in presenting them again here is not to make any specific policy recommendations, nor to commend our model or data above those of others, but to show how an innovative spirit and a model which provides global context to existing studies can give a new perspective and open up further avenues of research.

As we have tried to do throughout this thesis, we acknowledge that there are ways in which the model as it is presented here can be improved upon. We now critique our methodology in a concerted fashion.

11.3 A critique of the methodology

As the work of this thesis is based on input-output models, all the typical critiques of that methodology can also be levelled at our work. The primary of these are the use of linear production functions and a lack of an elasticity of substitution. It is our opinion that the second of these critiques is more justified than the first.

As to the first critique, we would argue that the use of arbitrary non-linearities in a model which seeks to be data-driven (viz. reflective of the "real world") has two detrimental effects. First, the structure of the global economy *as reflected by empirical observation* is distorted in ways which are by no means guaranteed to be an accurate reflection of a counter-factual. In other words, in the absence of knowledge to the contrary, the best assumptions are held by us to be those which distort the observed structure of the data as little as possible.

Secondly, our goal that model results be traceable, at least in theory, back to a set of empirical observations in the original data sets, is potentially derailed by the inclusion of arbitrary non-linearities. In other words, it becomes hard to say why the model results appear as they do.

Beyond the critiques of input-output modelling, which are well documented elsewhere, there are several possible critiques of our modelling approach which don't fall under the rubric of "recommendations for future work", covered in section 11.4. In our opinion, the most serious of these involves the decision to treat observed countries and estimated countries identically.

The 40 countries whose input-output tables (IOTs) are available in the World Input-Output Database (WIOD countries) are not the same as the 150 estimated countries whose IOTs come from the mathematical estimation analysed in chapter 5 and, arguably, should not be treated as such. Part of our motivation for the use of import ratios, as opposed to separate sets of technical coefficients for domestic and imported sector use, came from the desire to minimise the number of estimation points required to add each estimated country. Clearly, this goal of parsimony does not apply to the WIOD countries. It is therefore arguable that the IOTs of WIOD countries could have been used directly, without the simplifying assumption of import ratios.

Our counter-argument that this increases the complexity of the exposition of the model might be found insufficiently convincing for a sceptical audience. A further counter-argument that work on import substitution such as that presented in chapter 6 would be made more difficult without import ratios is perhaps slightly stronger.

WIOD countries and estimated countries could also have been treated differently in other more fundamental ways. An obvious example would be to use a reduced set of sectors for the IOTs of the estimated countries. An example of such a reduced set is given in table 5.1 on p. 144, as part of the discussion of chapter 5. This would, however, make far more complicated the exposition of results of the full model which covers both WIOD and estimated countries.

Furthermore, it is arguable that, for a model purporting to be data-driven, some of the analytical results are overly reliant on the estimation of point-to-point services flows from services totals as presented in chapter 4. The poor state of services flow data in the UN data set becomes increasingly problematic as the global economy becomes increasingly services-driven. It is, of course, far beyond the scope of this thesis to attempt to improve the state of that data, but perhaps our approach should be more explicit in distinguishing between commodities flows from data and services flows which have had to be estimated. The coincidence of poorly-gathered services data and the increasing reliance on services of many countries threatens the entire concept of a data-driven approach to global modelling. It would perhaps be more realistic to describe our model as being one of global *commodities* trade, rather than of global trade per se. It is up to the reader to decide whether or not the loss of the entire services side of the model would be justified by the improved adherence to the data-driven philosophy we have espoused throughout this thesis.

11.4 Recommendations for future work

Many of the critiques of the previous section can simply be viewed as opportunities to extend and improve the model presented here. Since our goal in creating it was at least partly as a proof of concept, showing that a data-driven approach to global modelling is both possible and useful, many of the possible additional complications have been left out as they would not have contributed to either of these efforts. But, with that caveat aside, there are several extensions which appear to us to be obvious next steps.

In our previous justifications of separation between the national and the international parts of the model (see section 3.4.4 on p. 98), we mentioned that a more complex view of global trade could be inserted in place of our extremely simplified, linear model. Depending on the particular persuasion of the model user, this may be viewed as the most pressing priority. A huge amount of work has been done on the determinants of trade (although the ability of this work to satisfactorily explain levels of global trade is debatable) and our assumptions, in the interests of simplicity, ignore this work completely.

On a different note, we mention several times throughout this work (section 8.1 on p. 206, section 10.3.2 on p. 268) that both WIOD and UN data are available as a time series and that, in theory, models can be analysed for each year for which we have data. In fact, such models *were* created, but the work of this thesis makes no attempt to use this extra dimension of analysis. An entire thesis-worth of additional analysis would be possible extending the analyses presented throughout part III of this work to include this time dimension. How have the results of each of these analyses changed over time? To what extent does the world look different after the 2008 financial crisis? Is there a pattern to the changes in, for example, technical coefficients, which would allow us to estimate input-output tables one period into the future?

Finally, during the creation of this thesis, several separate attempts were made at creating a web-based “front end” for the model, to allow users to create and explore scenarios without the need to understand the underlying Python code which drives the computer simulation of the model. At least one of these attempts was actually demonstrated at an event at the Royal Society in London, with users able to change the trade flows of a particular sector between a pair of countries, and watch the resulting changes in real time. However, none of this work was ever brought to completion and none of it features in any part of this thesis. A huge amount of developer time could be utilised by the revival of one or other of these projects, which would also allow our model to reach the policy-makers and non-technical researchers who are surely its natural audience.

These are just a few suggestions of how the model we have presented here might be used in future work. But clearly the very nature of the development of a new approach to modelling in any context means that the introductory analyses presented here are mere suggestions of the potential this approach opens up. Indeed, the particular specification of the model itself is only a suggestion of the ways in which our data-driven approach can and should be used by researchers to ask questions with global reach and ambition. The difficulty of working with the huge data sets involved, and the computing power required to solve such large systems of equations, can no longer remain an excuse for researchers to avoid analysing the global consequences, and causes, of the phenomena they study.

Appendices

Appendix A

Input-Output Tables

A.1 Description

As well as being used to calculate GDP figures, Input-output analysis is used in the economics literature to study supply chains, to avoid the double-counting of intermediate products in international trade statistics, Bems, Johnson, and Yi (2011) is a recent example, and also to look at how resources are allocated within a country (Jones 2011). But, generally, the economics literature has a tendency to overlook input-output in favour of less data-driven, more conceptually subtle and involved models (Basu (1995) is a notable exception).

Input-output analysis is, at its heart, nothing more than an accounting methodology. The economy of a city, region, country or even of the world, is divided into categories of economic activity, called sectors, based on the type of service provided or the type of product produced.

These sectors provide goods and services for consumption by the general public, and this is known as supplying to final demand. Examples include the selling of food to the customers of a restaurant, and the sale of cars direct to motorists. Consumption of goods by the government is also considered to be final demand.

But sectors also produce goods and services which are used as inputs to the production of goods and services in other industries. Input-output accounting records the extent to which the outputs of an industry are shared between final demand and intermediate use by each of the other industries.

A set of input-output accounts is presented as a matrix with the complete list of sectors both down the left-hand side and along the top. The cells of this matrix are then used to record the monetary value of the row sector's output which was purchased in a given year by the column's sector. Each number then represents an amount bought from one sector by another sector for intermediate use.

A.2 Use in modelling

When final demand is added to the matrix as a final column, the row sum represents the total amount of output, in monetary value, produced by the row sector in that year.

Since each number is an amount bought by the column sector, summing down the column will give the total amount spent on productive inputs in the given year. The relative sizes of these numbers, calculated by dividing each by the column total, then gives a sort of 'recipe' for producing a unit of the column sector's good. The recipe's elements are known as technical coefficients and describe the production technology of the column sector.

By assuming that these recipes for production are linear in their input factors, a simple form of dynamic analysis can be undertaken. For example, exogenously increasing consumer demand for a particular product, causes that sector's output to increase. The input-output table can be used to calculate the effect that this has on the sectors which supply intermediate products to the sector whose output has increased. In turn, the sectors which supply these sectors will also have to increase their output to meet the new demand for intermediate products which resulted from the exogenous increase in demand in final products. This will reverberate around the economy resulting in an increase in total output greater than output required to meet the exogenous increase in demand. This greater-than demanded increase in output is known as a multiplier effect and is one of the main benefits to using input-output analysis to describe an economy. A central agency whose goal is to maximise the total output effect of spending some additional money on a sector's output would choose the sector with the highest multiplier effect. These multiplier effects were first discussed by Leontief (1936) who gave a beautifully elegant matrix-algebra method of calculating them.

Multiplier effects assume that no industry is at or near capacity such that all industries are able to increase output to meet the exogenous change. In reality, the new demand may be met by an increase in imports, or by diverting exports towards domestic use. In these cases, the multiplier effect overstates the benefit to the economy.

Note that much of the above discussion is due to Miller and Blair (1985), who have an excellent descriptions of the types of analysis possible with input-output tables.

A.3 Global Multiregional Input-Output

When input-output tables, which describe the internal workings of an economy, are linked together using trade information, a multiregional input-output table (MRIO) is constructed. This MRIO is a description of the extent to which industries use intermediate goods from all other sectors, in all other regions. If the regions were national states, it would record, for example, how much of the output from the raw materials sector in China, the manufacturing sector in the US used in a given year.

The subject of global multiregional input-output (GMRIO) modelling has received considerable considerable attention recently, with an entire edition of the journal *Economic Systems Research* dedicated to new databases which attempt to build GMRIO tables. Much of the discussion which follows is expanded upon by Lenzen, Kanemoto, et al. (2012).

There are now at least four GMRIO databases, namely:

GTAP The Global Trade Analysis Project, an introduction to which is given by Walmsley, Aguiar, and Narayanan (2012) is possibly the first attempt at putting together the information which would be required to build a truely global MRIO table. It has data for the period 1992—2007, with approximately three-year periods within that range and the most recent release covers 129 countries with 57 sectors. It seeks to combine input-output tables with economic models to provide a framework for general equilibrium analysis.

EXIOBASE The EXIOBASE project, now concluded, was an EU project to create an environmentally-extended set of country accounts for 43 countries and 129 sectors. A complete description is given by Tukker, Koning, et al. (2013)

Eora A hugely ambitious GMRIO database covering 187 countries, with each country having differing numbers of sectors, depending on data availability and the perceived reliability of that data. Baseline IO tables are created for the year 2000, for countries which publish the relevant data. The IO tables for the other countries were then estimated using a baseline of an average between the Australian, Japanese and US tables, adjusted to fit available macroeconomic data. Other years in the time series were then estimated by iterating forwards and backwards according to observable time-series data.

WIOD For 40 countries and 35 sectors, a consistent time series in current and previous year prices of input-output tables and a world input-output table for each of the years 1995 to 2009. See below for more details, or refer to Timmer, Erumban, et al. (2012) for a complete description.

Appendix B

Data

B.1 Database engine

The data used in this thesis is relatively large in magnitude (somewhere in the region of 100GB) and arduous to obtain. (See the relevant *retrieval* sections of the proceeding subsections.) It was therefore important from the very start to have a reliable data system to ensure the data would remain available and uncorrupted. To this end, a dedicated Linux Red Hat server computer was purchased to which the project team had full administrator access. The database engine chosen had to be compatible with Linux, be capable of storing large quantities of data and be robust and stable. Since the data was potentially going to be of interest to the whole team, concurrency and user-level security was also important.

There are a number of open-source database engines which are available for free and are well supported by online communities. The choice between these database engines usually comes down to which client-side tools the developers in question are most familiar with. We now outline the most popular open-source database engines.

PostgreSQL owes its bizarre name to its history: it was developed as a successor to a database engine called “INteractive Graphics Retrieval System”or Ingres, hence “post-gres’. It prides itself on a commitment to the SQL standard, and on being a reliable, free enterprise-quality engine.

It is also well known for its GIS extensions, allowing for the storage and retrieval of spatial data in the form of co-ordinates for points, lines and polygons, with specifiable map projections. It is these GIS extensions which led to the choice of PostreSQL as the project database, although this thesis does not use them.

Known for its good built-in web-based administration tools, MySQL is the most popular open-source database engine. It has good user-level security, and concurrency, but does not implement a standardised version of the structured query language (SQL) which is used to interact with data. This lack of standardisation makes it harder to switch database engines in the future.

MySQL has very fast read performance, but is not as fast in concurrent read/write situations. MySQL is now owned by Oracle.

SQLite is a very light-weight minimal database engine. Each databases takes the form of a single file, rather than a suite of configuration files and running processes as with other engines.

This engine is used where simplicity, speed and low installation overheads are important such as for web applications and smartphones.

Appendix C

The world input-output database

C.1 Overview

The statistics authorities of many European countries and countries in the OECD produce national accounts in the form of supply and use tables. Each country has its own schedule for producing such national accounts, ranging from publication every year to twice in a decade.

Supply and use tables divide economic activity into industries, just as the input-output table does, but also into products. A supply table shows the extent to which the firms in a given industry produce each product and, often separately, use tables show how these products are used by the different industries. There have been various attempts to harmonise the systems for categorising both products and industries, each attempt with its own set of revisions over time, leading to a tangled web of standards and revisions in which supply and use data has been published.

It was in the context of this lack of temporally and internationally non-comparable national accounting systems that, in 2009, the European commission began funding for a project called the World Input-Output Database, or WIOD (Dietzenbacher et al. 2013). The project would bring together in one place a time series of comparable supply and use tables, bringing standards into line, adjusting to a consistent pricing system—supply and use tables sometimes report wholesale prices and sometimes consumer prices—and even filling in the gaps in the time series resulting from infrequent publication.

C.1.1 Constructing the data

Producing supply and use tables for years in which none were published was achieved by interpolating those supply and use tables which were available using an external time series for the country in question. This time series was often one of total output figures, but the details differed for each country in the data set. The whole procedure is described in detail, country by country, in Timmer, Erumban, et al. (2012), a document available from the WIOD website.

The project was also tasked with producing, from these standardised supply and use tables, a complete set of input-output tables, as described in the previous section, which are effectively a compaction and summary of the supply and use tables.

The project covered the years 1995 to 2009 and included 40 countries and represents a huge research effort in completing missing data using “other sources” and dealing with incompatibilities, missing data and oddities in each of the 40 countries.

As well as this effort, the WIOD project combined the standardised input-output tables they had produced with data published by the UN on international trade flows. COMTRADE (freely available for queries returning less than 50,000 rows at http://comtrade.un.org/db) is a one billion record database of the flow of goods between countries, with records going back to 1962. This product trade data, along with other data sources which attempt to cover trade in services (namely the OECD, WTO, Eurostat and the IMF) was used to produce world input-output tables, one for each of the years 1995-2009. These extremely ambitious tables—each year has over two million data points—combine the country input-output tables together showing explicitly where sector imports came from and where sector exports went to.

Finally, the WIOD project also gathered socio-economic accounts which record, among other variables, the amounts paid per sector for labour and capital, the number of employees per sector and the total hours worked by employees per sector. The labour compensation variables are divided into high-, medium- and low-skilled categories.

This represents an unprecedented amount of information, easily available, comparable between countries and interpolated for a continuous time series, about a relatively small number of countries which make up the vast majority—around 85 per cent—of the world’s total economic activity.

C.2 Conversion of WIOD to relational database form

The data available at wiod.org comes in the form of Excel spreadsheets. In particular, each country has a spreadsheet containing each of that country's NIOTs with each year's NIOT on a separate tab within the spreadsheet. Figure C.1 shows a schematic version of the structure of a NIOT as it is represented in the spreadsheets. All double-quoted text appears in the spreadsheet in exactly the form shown. The structure of the row/column headers is as follows.

C.2.1 NACE Codes

The top level of the header is a classification, relating the sector to a part of a classification system. Each of the 35 sectors is associated with a set of NACE[1] codes. NACE is a European Union system which groups economic activities (goods and services) into categories. The current version of NACE is revision 2, which was introduced in 2008, but the sectors of the WIOD are based on revision 1. NACE revision 1 has 17 top-level classifications, denoted by a single capital letter from A to Q. Some WIOD sectors map precisely to these top-level classifications. For example, the WIOD sector "Mining and Quarrying" maps precisely onto NACE code C.

Beneath these top-level classifications are several sub-levels. The first sub-level adds an additional letter to the first, and the second sub-level is denoted by a single number, which is unique throughout the classification. Thus the first top-level classification, A, contains a single letter sub-level, AA, which contains the numeric sub-levels 01 and 02. The first numeric sub-level of top-level classification B is then 05 (03 and 04 presumably being reserved for additional numeric sub-levels below A).

Some WIOD sectors map directly to one of the numeric sub-levels. Thus "Rubber and Plastics" maps directly to the numeric sub-level 25 which is labelled "Manufacture of rubber and plastic products" in the NACE classification itself.

Finally, some WIOD sectors map to a range of NACE classifications. In these instances, the range is indicated with the first classification, a lower-case letter t, then the last classification. For example, the WIOD sector "Textiles and Textile Products" covers classifications 17 and 18, "Manufacture of textiles" and "Manufacture of wearing apparel; dressing and dyeing of fur" respectively. This is represented by the string "17t18".

[1] From a French term whose English translation is "Statistical Classification of Economic Activities in the European Community"

[p]

				Nace codes Sector Name x 35	Final Demand & Investment x 5	“Exports”	“Total Output”
				Country	Country	Country	Country
				Column code	Column code	Column code	
NACE codes	Domestic input sectors x 35	Country	Column code	Domestic intermediates	Domestic FD	Domestic exports	Total domestic production
NACE codes	Imported input sectors x 35	“Imports”	Column code	Imported intermediates	Imported FD	0	0
	“Total Intermediate Consumption”	“TOT”	Row code				
	Assorted ‘other’ inputs x 4		Row code				
	“Value added at basic prices”		Row code				
	“International Transport Margins”						
	“Output at basic prices”		Row code				

Figure C.1: Structure of a national input-output table as downloaded in Excel spreadsheet format from wiod.org. FD is ‘final demand’.

Since, as we have seen, each WIOD sector maps to a number of NACE categories, there are a number of individual products making up each sector.

By categorising each product as either for intermediate use, for final consumption or for investment, WIOD makes assumptions about the final destination of all goods produced and imported in a given country. An assumption is made that the amount going to each destination is proportional to the extent to which each product category is present in the make-up of the sector in each country. It is to this classification which we now turn.

C.2.2 Use as intermediates

All use of economic output by corporations is considered intermediate consumption. The cell in figure C.1 labelled "Domestic intermediates" is a collapsed representation of what is actually a 35×35 matrix of cells in the original spreadsheet. Each element of this matrix represents the value (in $US millions) of a flow from one domestic sector (depending on the *row* in question) to another (depending on the *column* in question). These goods or services are used as intermediates in the receiving sector's production process.

Similarly, the cell labelled "Imported intermediates" represents a matrix of the same dimensions, but representing flows of intermediate goods originating abroad.

Together, these two blocks make up the Z matrix of a classic input-output table, and are used in the calculation of technical coefficients for this country in this year.

C.2.3 Final demand and investment

In contrast to intermediate consumption, households, non-profit institutions and government have final consumption, and each of these is considered a sector in the NIOT. The light-blue "final demand and investment" columns represent flows from sectors, both domestic and imported as above, to these final consumption categories. The five columns are[2]:

Final consumption expenditure by households both private households and "institutional households" such as hospitals, residential homes, prisons etc.

NPISH Final consumption expenditure by non-profit organisations serving households. Non-governmental agencies which provide goods or services to households at an economically insignificant price (or for free). This includes religious societies and sports societies.

---

[2] See the eurostat glossary at epp.eurostat.ec.europa.eu for more detailed information.

Final consumption expenditure by government includes only goods and services purchased for the production of non-market final goods and services, or for goods and services provided as social transfers[3]

Gross fixed capital formation defined as the investment in tangible and non-tangible assets by domestic producers, deducting disposals.

Changes in inventories and valuables the difference between additions to and withdrawals from inventories plus the acquisition of valuables such as precious metals, antiques and works of art.e

The first three of these are final demand columns, and the last two are investment columns. The goods and services flowing to any of these columns are assumed to be completely consumed within the country and in the year in question.

C.2.4 Exports

The "Exports" column shows, in the first 35-element block, total exports of domestic goods. This, together with the uppermost intermediate flows block and the uppermost final demand block, defines all possible destinations for domestically produced goods: they are either used as intermediates, consumed by final demand, or exported.

The difference is that no imported goods are allowed to be directly exported. (Although they clearly *can* be exported as part of goods which have been worked on domestically.)

The final column, "Total output" is simply a sum of the various destinations outlined above, for domestically produced goods only; the column contains zeros for the rows relating to import sectors.

Together with the 35 intermediate flow columns, the five final demand and investment columns and the export and total columns make each NIOT 42 rows wide.

C.2.5 Additional rows

Beyond the 35 domestic sector rows, and the 35 imported sector rows, each NIOT has a number of additional rows.

The first sky-blue row, labelled "Total Intermediate Consumption", is a column sum of the "Domestic intermediates" and "Imported intermediates" blocks. There then follow four columns, in salmon pink, representing assorted "other" inputs in either the production process or to final demand.

[3] See www.oecd-ilibrary.org for full details.

The first of these is labelled “taxes less subsidies on products” and represents the amount the “recipient” industry, i.e. the industry corresponding to the column in question, had to pay in the form of taxes, net government subsidies. In heavily subsidised industries, such as Agriculture, this can be negative.

The second row relates to differences in the way that transport costs are represented between total imports and the import figures for single goods. Total imports are measured “free on board” (FOB), meaning that the transport costs are ignored from the perspective of the importing country. But the import of a particular product can be measured with “cost, insurance, freight” (CIF). The row “Cif/ fob adjustments on exports” is way to account for differences in reporting method between these two levels of reporting. WIOD contains no non-zero values in this row.

The next two additional rows both relate to final demand columns. “Direct purchases abroad by residents” tracks the contribution to final demand of goods which were produced abroad, and bought abroad without being imported into the country in question. This row is only non-zero in the household and government consumption columns (and almost exclusively in the household column.) “Purchases on the domestic territory by non-residents” counts the money spent domestically by people from abroad. This can sometimes be negative, as a result of the balancing procedures used to create the WIOD tables in the first instance.

“International Transport Margins” is a cost related to transport services provided by foreign labour and capital. A typical example of this is the payment of shipping costs a domestic sector makes when the ship or its crew is owned by a third-party country. Streicher and Stehrer (2012) give a more detailed explanation of how and why such costs are incurred.

The row “value added at basic prices” is an accounting identity. It is equal to the “Output at basic prices” minus “Total intermediate consumption” and all other salmon pink adjustment rows outlined above. This is an important quantity, as it described the difference between a sector’s output and all of its various inputs and costs.

Table C.1: Relational database format for storing input-output flow data

ID	Country (ISO3)	Year	From Sector	To Sector	Flow Value
1	AUS	1995	1	1	3107
2	AUS	1995	1	2	12
3	AUS	1995	1	3	8043

C.3 The structure of the input-output tables in relational form

As mentioned above, each country in WIOD has its own spreadsheet for NIOTs, and each year's NIOT is on a separate tab. In order to make the data as easily available as possible, we wanted the data in a single table in a relational database. The standard representation of data in a relational database treats each data point as an observation of some real world entity. The 'entities' in an input-output table are sector flows: a flow of value, in the form of goods and services, from one sector to anther. Here, we use the term "sector" to mean any source or destination of value, including intermediate demand, final demand and exports.

Table C.1 shows the structure of an NIOT as stored in the Postgres relational database. Notice the following important elements of each row. Each row has a unique identifier called "sector_flow_id". The country is stored as an ISO 3166 3-letter country code[4]. The sector names are replaced by IDs, which reference a separate table "sectors" which contains the names of each of the WIOD sectors along with a unique identifier and some metadata which facilitates retrieval, as explained in section C.3.1.

Each row of the sectors table has a unique identifier, the name as it appears on the WIOD spreadsheets, the NACE codes as they appear on the WIOD spreadsheets and a short name which is used in all analysis. In addition to these fields, each row has five meta-data flags to assist with filtering the data during retrieval and analysis. The complete list of which sectors are in which category is shown in table C.3. The most important of these categories are "is_final_demand" and "is_investment". The columns in these categories are summed to create a single final demand and single investment column, respectively.

C.3.1 Procedure for uploading Excel data to Postgres

In order to convert the data in the WIOD spreadsheet format shown in figure C.1 to the relational-database format discussed above, significant processing of the data is

[4] see http://www.iso.org/iso/country_codes.htm for details

required. This involved writing a script in Visual Basic for Applications (VBA, the scripting language which lies behind Microsoft Office products such as Excel) to automate the process as far as possible. It is important that the process be as automatic as possible to prepare for the possibility that the WIOD team produces additional datasets, either for more countries or for more years.

In brief, the VBA script moves through each input-output table left-to-right then top-to-bottom, recording the value of the flow in the current cell and also the row and column the value came from. The script then uses this information first to look up the sector ID (or create a new row for the sector if it doesn't yet exist in the sectors table) for both the source and the destination sectors, then to build a SQL INSERT query of the form:

```
INSERT INTO sectors(country_iso3, year,
   from_sector_id, to_sector_id, flow_amount)
VALUES ('AUS', 1995, 1, 1, 5693);
```

Given a list of Excel file names, and a list of tab names contained in each Excel file (where tab names define which year a particular NIOT relates to), the VBA script will then run automatically through each Excel file in the list, and each tab in each Excel file, and generate these insert queries, one for each element of each NIOT.

C.4 Sectors

Here, in table C.2, we give a complete description of the sector names, along with their NACE codes and, crucially, the short name used to refer to the sector everywhere in this thesis. We also outline in table C.3 how each sector is categorised using a system of on/off flags for ease of retrieval.

Table C.2: The WIOD sectors as stored in the Postgres relational database

ID	Name	NACE	Short Name
1	Agriculture, Hunting, Forestry and Fishing	AtB	Agriculture
2	Mining and Quarrying	C	Mining
3	Food, Beverages and Tobacco	15t16	Food
4	Textiles and Textile Products	17t18	Textiles
5	Leather, Leather and Footwear	19	Leather
6	Wood and Products of Wood and Cork	20	Wood
7	Pulp, Paper, Printing and Publishing	21t22	Paper
8	Coke, Refined Petroleum and Nuclear Fuel	23	Fuel
9	Chemicals and Chemical Products	24	Chemicals
10	Rubber and Plastics	25	Plastics
11	Other Non-Metallic Mineral	26	Minerals
12	Basic Metals and Fabricated Metal	27t28	Metals
13	Machinery, Nec	29	Machinery
14	Electrical and Optical Equipment	30t33	Electricals
15	Transport Equipment	34t35	Vehicles
16	Manufacturing, Nec & Recycling	36t37	Manufacturing
17	Electricity, Gas and Water Supply	E	Utilities
18	Construction	F	Construction
19	Sale, Maintenance and Repair of Motor Vehicles and Motorcycles & Retail Sale of Fuel	50	Vehicle Trade
20	Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles	51	Wholesale Trade
21	Retail Trade, Except of Motor Vehicles and Motorcycles & Repair of Household Goods	52	Retail Trade
22	Hotels and Restaurants	H	Hospitality
23	Inland Transport	60	Inland Transport
24	Water Transport	61	Water Transport
25	Air Transport	62	Air Transport
26	Other Supporting and Auxiliary Transport Activities & Activities of Travel Agencies	63	Transport Services
27	Post and Telecommunications	64	Communications
28	Financial Intermediation	J	Financial Services
29	Real Estate Activities	70	Real Estate
30	Renting of M&Eq and Other Business Activities	71t74	Business Services
31	Public Admin and Defence & Compulsory Social Security	L	Public Services
32	Education	M	Education
33	Health and Social Work	N	Health
34	Other Community, Social and Personal Services	O	Other Services
35	Private Households with Employed Persons	P	Private Households
36	Final consumption expenditure by households		Household Demand

Table C.2: (continued)

ID	Name	NACE	Short Name
37	Final consumption expenditure by non-profit organisations serving households (NPISH)		Non-profit Demand
38	Final consumption expenditure by government		Govt Demand
39	Gross fixed capital formation		Savings
40	Changes in inventories and valuables		Inventory
41	Exports		Exports
79	taxes less subsidies on products		Taxes
80	Cif/ fob adjustments on exports		Export Adjustments
81	Direct purchases abroad by residents		Overseas Purchases
82	Purchases on the domestic territory by non-residents		Visitor Purchases
83	Value added at basic prices		Value Added
84	International Transport Margins		International Transport Margins
126	Final Demand (combined)		Final Demand
127	Investment (combined)		Investment

Table C.3: The metadata flags of the sectors table. Sectors listed have this flag set to True

Metadata flag	Sectors
is_production_sector	1–35
is_final_demand	36, 37, 38
is_investment	39, 40
is_export	41
is_value_added	79, 80, 81, 82, 83, 84

C.5 Retrieving the input-output data

We use database views and a Python data-handling library called Pandas[5] to retrieve the data from Postgres and reconstruct the elements of a NIOT. First, the `sector_flows` table is joined to the `sectors` table twice, once for the source sector and once for the destination:

```
SELECT * FROM sector_flows sf
INNER JOIN sectors s1 ON sf.from_sector_id = s1.
   sector_id
INNER JOIN sectors s2 ON sf.to_sector_id = s2.
   sector_id
```

This allows us to use the short sector names from table C.2 in the analysis which follows. We then use the metadata flags as described in table C.3 to select only those rows which are required for each part of the table.

For example, when reconstructing the 35×35-element block at the top-left of the table (the intermediate flow section) we append:

```
WHERE is_production_sector
```

to the above query. To reconstruct the 35-element final demand vector we append:

```
WHERE is_final_demand
```

The results of these queries are saved as comma-separated variables (CSV) files for import into Python using the Pandas library mentioned above.

To restore the 'wide-format' style of an input-output table, the Pandas function `unstack` can be used. This function takes one column of a data table as a source column, and generates a new column for each distinct value the source column takes, pivotting the data into wide format. Therefore by "unstacking" the column denoting the recipient sector, a square 35×35 matrix is reconstructed from the long-format database output.

Finally, to create a single final demand vector from the three sectors which have `is_final_demand`, the values are summed across the sectors. The same is done for the two sectors which have `is_investment` in creating a single investment sector.

[5] see http://pandas.pydata.org/ for a complete description of the package

Appendix D

UN COMTRADE

D.1 Overview

In this section, we discuss the data source used to create the between-country trade aspect of the model. The database is called the United Nations commodity trade database, or COMTRADE. We first give some background about the database itself, then discuss the data, its retrieval and storage, and finally give some comments on handling large datasets.

D.1.1 Background of the COMTRADE database

COMTRADE is the United Nations' database of commodities trade, based on member countries' customs reports about the quantity and value of physical commodities leaving and arriving from their borders. The data is gathered by the countries for tax purposes and therefore has excellent coverage, since the countries have an incentive themselves to gather good data. The data is then reported to the UN and made available via the UN website at comtrade.un.org.

D.1.2 The choice of COMTRADE and possible alternatives

Although COMTRADE has unparalleled coverage, it is less easily available as an entire dataset than other more coarse-grained sources.

Many papers on trade use Feenstra et al. (2005) which has coverage from 1962 to 2000. It provides trade quantities and values at the SITC[1] 4-digit level for around 50 countries. The data is in easily-downloaded STATA format which explains its huge popularity. It values importer-reported data over exporter-reported wherever possible.

[1] Standard International Trade Classification

Another very popular trade dataset is that from the *Correlates of War* (CoW) project[2]. This has dyadic (i.e. importer-exporter pair totals) data from 1870 to 2009. It is largely based on the IMF's *Direction of Trade Statistics*[3] database of dyadic totals.

[2] http://www.correlatesofwar.org
[3] http://elibrary-data.imf.org/

Table D.1: Table showing COMTRADE trade flow records for 2011 for commodity code H4-961310 "*Pocket lighters, gas fuelled, non-refillable* between France and the UK

Trade Flow	Reporter	Partner	Qty	Weight (kg)	Trade Value (US$)
Import	France	France	21,602	1,104	6,8170
Re-Import	France	France	21,602	1,104	6,8170
Import	France	UK	2,409	277	5,342
Export	France	UK	494,773	18,316	748,229
Import	UK	France	386,841	11,218	607,532
Export	UK	France	37,627	2,899	36,219

D.2 Data which COMTRADE makes available

COMTRADE provides data on trade flows of physical commodities at the six-figure Harmonized System (HS) classification. This is an extremely detailed level, with around 5,000 distinct classifications in the most recent revision (Revision 3, with codes starting "H3"). As an example of the level of granularity, there are four categories relating to cigarette lighters: refillable, non-refillable, parts thereof and "*Lighters other than pocket lighters, whether/not mechanical/electrical*".

Table D.1 shows the raw data as it comes directly from the COMTRADE website (although several identifier columns and descriptive columns have been removed for display purposes.) Rather than a *from country to country* format, COMTRADE reports "Reporter" country and "Partner" country. Then, depending on the direction of the flow (according to the column "Trade Flow"), both countries can act either as importer or exporter for any given record.

D.2.1 Countries versus trade areas

COMTRADE contains reports of trade flows to and from 288 distinct trade areas of which 258 have ISO3 country codes (three-letter country codes which uniquely identify a country; GBR is Great Britain, FRA is France, etc.). The remaining 30 are regions used as trade "partners" when the origin or destination of goods was not known or incompletely recorded by the reporting authority. Examples of this include "Europe EU, not elsewhere specified", "Western Asia, not elsewhere specified" and even the generic catch-all "Areas, not elsewhere specified". Flows to these areas without ISO3 code were all assigned to ISO3 code "XXX". In this way, calculations of trade totals by summing individual flows can still be consistent with the original data without the need to invent a series of ISO3 codes for regions.

Of the 258 which *do* have country codes, 24 are listed as having ended before the date range of the present work. (1995 to 2011, as defined by the date range of WIOD, see above.) Examples include "Czechoslovakia", "Réunion" (which was later classified as being part of France), and "Former Democratic Republic of Germany". Just three ended between 1995 and the present day: Belgium-Luxembourg, which became separate trade areas in 1998, the "Southern African Customs Union" (ZAF) which ended in 1999 and Serbia and Montenegro which stopped reporting jointly in 2005.

Within the remaining countries, there is still a huge grey area between countries and regions, largely because the trade data is captured at whichever level of granularity was most convenient for the reporting authorities, rather than as a standard list of agreed countries. Thus both Hong Kong and Macao are listed as separate entities with their own ISO3 codes, as are the Cayman and Christmas and Falkland Islands. At the other end of the granularity scale is a "World" (WLD) trade region for recording flows where either the origin or destination was completely unrecorded.

Flows to and from these regions are dealt with by the introduction of the concept of a "Rest of the World" (RoW) which acts throughout this work similarly to a normal country, and serves as a sink and source for flows involving trade areas which would not normally be described as countries.

D.2.2 Classification systems and their revisions

With the compilation of data reported by 200 different trade areas, and over a long time period, comes the inevitable problem of different classification standards. There are many international classification standards for commodities, of which NACE, used by the World Input-Output Database (see section C.2) is just one.

COMTRADE offers data classified by two systems, the Standard International Trade Classification (SITC) and the Harmonized System (HS). The former is published and revised by the UN, and the latter by the World Customs Organisation[4]. The HS is used most often for data reporting, and the SITC is a less specific categorisation for analysis purposes[5].

The HS is currently on its 5th revision with each revision making a relatively small change to the categorisation[6]. Each revision has a two-digit code representing the year in which is occured, and also a revision number (starting at zero). For our purposes, the differences between revisions are entirely assumed away, with 6-digits codes being

---

4 www.wcoomd.org/

5 For details, see http://bit.ly/25BcrUO

6 For full details see http://www.globaltariff.com/HSHistory.cfm

assumed at least to fit into the same economic sectors between revisions. This seems an innocuous assumption since most sectors are fully defined far above the six-figure level.

By way of an example, the 3rd revision of HS contains just over 5,000 product categories at the six-figure level.

D.3 Retrieval of COMTRADE data from the web interface

At the time the data was compiled, the only way to retrieve data from COMTRADE was via requests submitted to the COMTRADE website[7]. This has since been replaced by a far more user-friendly API, allowing a more automated approach to the retrieval of data. The approach outlined here predates the existence of this facility.

D.3.1 Requesting data and request limits

We had the goal of completely reproducing the COMTRADE database for the period 1995-2011 which is the same as the period covered by WIOD.

Data can be requested by means of a query submitted through a form on the COMTRADE website. The data can be filtered by year, reporter country, partner country (the 'other' country involved in the flow, either the origin or the destination depending on the type of record), product code (down to the 6-figure HS level) and direction (import, export, re-import, re-export.) For the purposes of this study, we ignore re-imports and re-exports and focus simply on imports and exports.

Queries using these filters are submitted to the website, and will be processed by the server. When the server is ready to deliver the results set (in a period which can range from a few seconds to a few days, depending on the size of the query and the load of the server), the user is notified by email.

Submitted queries are listed on a page of the website called the "Batch Monitor". This page tracks the status of all submitted queries, either as "Submitted", "Inserted" which implies the query is running, and "Ready". When the status is set to "Ready" the Batch Monitor will provide a link for the direct download of a CSV file of results, archived into a .zip file.

Users are limited to the size of dataset a single query can return. For standard users this limit is 50,000 records. For a fee of US$1,000 this limit can be increased to 5,000,000. At the time of submitting the query, the server does not know whether this limit will be exceeded or not. Thus, any query can be submitted and, once the meta-data for the query has been calculated, will either be accepted by the system or rejected. This process can take up to hours to complete. It is therefore important to formulate queries such that the row limit is not reached, otherwise the query will need to be submitted

[7] comtrade.un.org/db

again. Equally, the number of requests should be kept to a minimum to manage the complexity of the task of retrieving all records from the database.

To ensure that the query limit was not hit, the requests were built as follows: each country was submitted as a separate request with all import and export records from the years 1995 to 2011, in every 6-digit product category. This seemed a natural way to split the requests. It allowed us to make as few requests as possible (only as many as the number of trade areas, around 200) while minimising the possibility that a single request would reach the 5,000,000 record limit[8].

D.3.2 Known data errors

Some trade flows are recorded as negative. Of these, a portion are the 16-bit long integer underflow value -2,147,483,648. These are impossible to deal with as they clearly contain no information at all. There is a large range of other, sometimes very large, negative values and these are all included "as is" in everything in this thesis. No attempt has been made to correct for any of these negative values.

[8] The maximum theoretical size of one of these requests occurs if a country trades every product with every trade area in every year. In this case, $5{,}000 \times 200 \times 15 = 15{,}000{,}000$ records would be returned. While this would indeed exceed the number of requests, we thought this maximum to be *extremely* unlikely to occur.

Appendix E

The philosophy behind our modelling approach

In this appendix we present a blog, written by the author and posted online in August 2013, before much of the work of this thesis was completed. It is too informal for the main body of the thesis, but we feel it provides an interesting view of the "philosophy" which has driven us in producing the model and the analyses presented here.

> A recent meeting of minds at MIT[1], described as an "anti-disciplinary conference for the intellectually promiscuous," can have left little doubt in the minds of its attendees: the network is the paradigm *du jour*. The whole event was awash with the edges and nodes of an on-trend way of thinking about the world.
>
> The social sciences are buzzing with excitement about network analysis: here at UCL's Centre for Advanced Spatial Analysis (CASA), networks are being used in ways both intuitive, in the modelling of traffic flows around road networks and patterns in the use of social networking; and highly counter-intuitive, such as in the study of where and when crimes occur in relation to one another, how rioters and the police play out a networked game of cat and mouse, or how ethnic conflicts are affected by access to, and use by the government of, both new and traditional media.

[1] This was the "Links 2013" conference on the subject of economic complexity.

But economics has been late to the game. Eric Beinhocker has written forcefully about the need for economics to adapt, by doing nothing less controversial than simply to acknowledge the existence of the Second Law of Thermodynamics—pithily summarised by Jagger/Richards as 'you can't always get what you want'—and to embrace networks and the science associated with them: complexity.

Here are some examples of how this has *not* been happening in economics. The cost at which a particular government can borrow from investors is clearly related to how convinced the money markets are that other governments represent a safe bet, yet borrowing costs are routinely studied using only the variables of the country in question. Similarly, you can't expect to explain movements in macroeconomic variables such as exchange rates or foreign trade, without including the other countries in your analysis. It is this kind of old-world thinking, regress A against B to see if it has `an effect', that has led Economics to consistently misunderstand the world as it really is today. (And, indeed, as it always has been.)

John Stuart Mill, in motivating previous generations of economists, described the booms and busts of macroeconomics as being like a stormy sea tossing and rolling a boat:

"Would you advise those who go to sea to deny the wind and the waves — or to make use of them and to find the means of guarding against their dangers?"

This is an analogy which seems immediately to require a global focus to the study of Economics, and an approach to the systems and interactions involved which is meteorological in scale. But, to extend the shipping analogy, we too often see attempts to predict the movement of the ship based on the actions of the crew, rather than an acknowledgement that bigger forces are at work.

So to the problem of applying the lessons of complexity to the global economic system: we must start with a network upon which to operate. But the systems involved are intimidatingly large. The European Union, the globe's biggest single market, represents the economic activity of some 500 million inhabitants trading goods and services with a value of around $16 trillion. Any practicable representation of a network of systems of this complexity (and the global economy is surely the system of maximum-conceivable complexity) would be such a gross simplification that it would iron out all of the interactions, chain reactions, bifurcations and time-lag effects which could potentially make a model useful for explaining or forecasting global events. And yet we must begin somewhere.

As with the meteorologist, the place to begin must surely be with data. Given the dismal state of Economics as a tool for predicting even the simplest of human interactions, let alone something as subtle and sensitive as the famous but little understood "market outcome", we need inductive models based on as much data as we can gather. (An inductive model is one which is driven, to the greatest extent possible, by what can actually be observed, as opposed to theorised about.)

So how can a network representation of the global economy be put together in a way which is rich in data and light on assumptions? Perhaps we might continue the meteorological analogy: just as the weather is something greater than the movement of warm and cold air around the world, and yet can be described and predicted by knowing about how, where and when the air moves, so the economy is a function of, but something more than, the global flow of goods and services. And yet the phenomena of interest, wealth and growth, poverty and inequality, might be studied by watching these goods and services move around the world as the storms and heatwaves are predicted by weathermen watching airflows. Note that there's no assumption about causality in weather forecasting, or even any statement about the mechanisms by which certain air flows result in certain weather patterns. This is why an inductive model like this is distinct from a deductive one, where theories are postulated to explain mechanisms and causal links.

The flow of goods and services around the global economy can be described as taking place at two distinct levels: within an economy, and between economies. These two levels can each be characterised by a simple question.

Within an economy:
To what extent does the production of goods and services, the level of which is set in response to demand from consumers, require the production of other goods and services as inputs to the production processes? The answer to this question can be represented by a network of dependency between the various goods and services an economy produces. Some goods require a great deal of input from other sectors of the economy, others are produced with almost no input at all.

Between economies:
To what extent does trade in goods and services occur between economies which produce a product and economies which consume it? The answer again comes in the form of a network of trading relationships. Some countries have a history of trading certain products with one another, and other countries are largely self-sufficient. Critically, in both cases, the structure of the network can be derived purely from data.

In both cases, we build a network representation of flows which take the available data as a starting point, and take that data seriously. Where traditional economic models attempt to explain the high levels of trade between, say, France and its former colonies, here we let the data do the talking. If France trades more with Senegal than it does with Ethiopia, a former Italian colony, then the data will reflect this, and the network which is built on that data will include the bias in its construction, along with all other biases too numerous to hope to describe explicitly. We are taking the observable state of the trade network as the best possible proxy for all the subtleties and irrationalities of human interaction and, in doing so, explicitly abandon the attempt to include such unknowables in our model. It is in this sense that we are taking the data seriously.

Even with nothing but a static description of the world's economy as viewed through the observed networks of production and trade, we can perform some interesting analysis. For example, which trade link—that is to say, which product, traded between which two countries—contributes most to global output? The answer may not be simply the trade link with the highest dollar value. For example, Belgium provides around a quarter of France's imported iron and steel. If those two countries had a falling out which caused them to stop trading altogether then, all things being equal, France would have 25% less steel to use in manufacturing other goods. The French car industry uses 75% domestic metals and 25% imported metal, meaning that French output of cars would be down 6.25%. Since 50% of all cars produced in France are exported, and France exports over $20 billion-worth of cars a year, the spat between Belgium and France would cost the world over $600 million in lost car exports alone.

The above simple example is constructed using trade data from the UN, and a sector-by-sector description of France's economy produced by the European Union. By combining these two data sets we build a picture of the movement of goods and services not only through economies but between economies and, hence, around the world. By adding some simple linearity assumptions—to make 1% more cars, France needs 1% more steel—we can even begin to do some dynamic analysis of the global economy, to answer questions such as: how will the world respond to a growth in demand for cars coming from China? Or: which products would most benefit Nigeria if imports were replaced with domestic production? And how would that affect Nigeria's trading partners?

We describe this framework, a network of networks based on observable data, as the demonstration model: it is a least-assumptions description of the world's economy designed to give a demonstration of the power of network and complexity science when applied to questions which are unavoidably global in nature, such as those around migration, trade, international security and development aid.

Here at CASA we're currently assembling the data that will allow us to build our global network of networks. We're expecting to see the first version of our model, and the first of the demonstration analyses based on it, to be completed in the autumn.

Bibliography

Abel, Guy J. (2010). "Estimation of International Migration Flow Tables in Europe". *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 173.4, pp. 797–825.

Abel, Guy J. and Nikola Sander (Mar. 28, 2014). "Quantifying Global International Migration Flows". *Science* 343, pp. 1520–1522.

Acemoglu, Daron and Jaume Ventura (Jan. 5, 2002). "The World Income Distribution". *The Quarterly Journal of Economics* 117.2, pp. 659–694.

Adelman, Irma and Sherman Robinson (1978). *Income Distrtibution Policy in Developing Countries: A Case Study of Korea*. Oxford ; New York: Published for the World Bank [by] Oxford University Press. 346 pp.

Aitken, Norman D. (1973). "The Effect of the EEC and EFTA on European Trade: A Temporal Cross-Section Analysis". *The American Economic Review* 63.5, pp. 881–892.

Aizenmann, Joshua and Reuven Glick (2006). "Military Expenditure, Threats, and Growth". *The Journal of International Trade and Development* 15.2, pp. 129–155.

Alptekin, Aynur and Paul Levine (Dec. 2012). "Military Expenditure and Economic Growth: A Meta-Analysis". *European Journal of Political Economy* 28.4, pp. 636–650.

Anderson, James E. (Dec. 2010). *The Gravity Model*. Working Paper 16576. National Bureau of Economic Research.

Anderson, James E. and Eric van Wincoop (Mar. 1, 2003). "Gravity with Gravitas: A Solution to the Border Puzzle". *The American Economic Review* 93.1, pp. 170–192.

Armington, Paul S. (1969a). "A Theory of Demand for Products Distinguished by Place of Production". *Staff Papers (International Monetary Fund)* 16.1, pp. 159–178.

Armington, Paul S. (1969b). "The Geographic Pattern of Trade and the Effects of Price Changes (Structure Géographique Des Échanges et Incidences Des Variations de Prix) (Estructura Geográfica Del Comercio y Efectos de La Variación de Los Precios)". *Staff Papers (International Monetary Fund)* 16.2, pp. 179–201.

Askegaard, Søren, Eric J. Arnould, and Dannie Kjeldgaard (June 1, 2005). “Postassimilationist Ethnic Consumer Research: Qualifications and Extensions”. *Journal of Consumer Research* 32.1, pp. 160–170.

Athanassiou, Emmanuel, Christos Kollias, and Stavros Zografakis (Jan. 1, 2002). “The Effects of Defence Spending Reductions: A CGE Estimation of the Foregone Peace Dividend in the Case of Greece”. *Defence and Peace Economics* 13.2, pp. 109–119.

Baier, Scott L. and Jeffrey H. Bergstrand (Mar. 2007). “Do free trade agreements actually increase members’ international trade?” *Journal of International Economics* 71.1, pp. 72–95.

— (Feb. 1, 2009). “Bonus Vetus OLS: A Simple Method for Approximating International Trade-Cost Effects Using the Gravity Equation”. *Journal of International Economics* 77.1, pp. 77–85.

Balassa, Bela (1967). “Trade Creation and Trade Diversion in the European Common Market”. *The Economic Journal* 77.305, pp. 1–21.

Baldwin, Richard and Daria Taglioni (Sept. 2006). *Gravity for Dummies and Dummies for Gravity Equations*. Working Paper 12516. National Bureau of Economic Research.

Baldwin, Robert E. (2000). “Trade and Growth: Still disagreement about the relationships”. *Working papers - Organisation for Economic Cooperation and Development Economics Department*.

Ball, R. J. (1967). “Review of An Econometric Study of International Trade Flows”. *The Economic Journal* 77.306. In collab. with H. Linnemann, pp. 366–368.

— (1973). *The International Linkage of National Economic Models*. North-Holland. 488 pp.

Bangake, Chrysost and Jude Eggoh (Jan. 1, 2009). “The impace of currentcy unions on trade: lessons from CFA Franc Zone and implications for proposed African monetary unions”. *Savings and Development* 33.1, pp. 61–72.

Barro, Robert J. (Oct. 1990). “Government Spending in a Simple Model of Endogeneous Growth”. *Journal of Political Economy* 98 (5, Part 2), S103–S125.

Barro, Robert J. and Xavier Sala-i-Martin (1997). “Technological Diffusion, Convergence, and Growth”. *Journal of Economic growth* 2.1, pp. 1–26.

Basu, Susanto (June 1, 1995). “Intermediate Goods and Business Cycles: Implications for Productivity and Welfare”. *The American Economic Review* 85.3, pp. 512–531.

Baudains, Peter (2016). “Richardson Models with Space”. *Approaches to Geo-Mathematical Modelling: New Tools for Complexity Science*. Ed. by Alan G. Wilson. John Wiley & Sons.

Baudains, Peter, Hannah M. Fry, Toby P. Davies, and Alan G. Wilson (2016). “A Dynamic Spatial Model of Conflict Escalation”. *European Journal of Applied Mathematics* 27.3, pp. 530–553.

Baudains, Peter and Alan G. Wilson (2016). “Conflict Modelling: Spatial Interaction as Threat”. *Global Dynamics: Approaches from Complexity Science*. Ed. by Alan Wilson. First Edition. John Wiley & Sons, pp. 145–158.

Beaumont, Paul, Ingmar Prucha, and Vladimir Filatov (1979). “Performance of the LINK System: 1970 versus 1975 Base Year Trade Share Matrix”. *Empirical Economics* 4.1, pp. 11–41.

Becker, Gary S. (1952). “A Note on Multi-Country Trade”. *The American Economic Review* 42.4, pp. 558–568.

Beckerman, W. (1956). “The World Trade Multiplier and the Stability of World Trade, 1938 to 1953”. *Econometrica* 24.3, pp. 239–252.

Behar, Alberto, Philip Manners, and Benjamin D. Nelson (Dec. 2013). “Exports and International Logistics”. *Oxford Bulletin of Economics and Statistics* 75.6, pp. 855–886.

Beine, Michel, Frédéric Docquier, and Çağlar Özden (May 2011). “Diasporas”. *Journal of Development Economics*. Symposium on Globalization and Brain Drain 95.1, pp. 30–41.

Beinhocker, Eric D. (June 1, 2006). *Origin of Wealth: Evolution, Complexity, and the Radical Remaking of Economics*. 1st ed. Harvard Business School Press. 527 pp.

Bems, Rudolfs, Robert C Johnson, and Kei-Mu Yi (May 2011). “Vertical Linkages and the Collapse of Global Trade”. *American Economic Review* 101.3, pp. 308–312.

Bénard, Jean (1963). “Réseau Des Échanges Internationaux et Planification Ouverte”. *Actualité Économique* 39.3, p. 537.

Bergstrand, Jeffrey H. (Feb. 1989). “The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade”. *The Review of Economics and Statistics* 71.1, p. 143.

Berry, Brian (1966). *Essays on Commodity Flows and the Spatial Structure of the Indian Economy*. Department of Geography, University of Chicago.

Bertoli, Simone and Jesús Fernández-Huertas Moraga (May 2013). “Multilateral resistance to migration”. *Journal of Development Economics*. Migration and Development 102, pp. 79–100.

Biswas, Basudeb and Rati Ram (1986). “Military Expenditures and Economic Growth in Less Developed Countries: An Augmented Model and Further Evidence”. *Economic Development and Cultural Change* 34.2, pp. 361–372.

Bjerkholt, Olav (2009). “The Making of the Leif Johansen Multi-Sectoral Model”. *History of Economic Ideas*, pp. 103–126.

Blondel, Vincent D, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre (Oct. 9, 2008). "Fast unfolding of communities in large networks". *Journal of Statistical Mechanics: Theory and Experiment* 2008.10, P10008.

Bloom, David E., Jeffrey D. Sachs, Paul Collier, and Christopher Udry (1998). "Geography, Demography, and Economic Growth in Africa". *Brookings Papers on Economic Activity* 1998.2, pp. 207–295.

Böhmelt, Tobias and Vincenzo Bove (2014). "Forecasting Military Expenditure". *Research and Politics* 1, pp. 1–8.

Brockmeier, Martina (1996). "A Graphical Exposition of the GTAP Model". GTAP Technical Papers 5.

Bruton, Henry (1989). "Chapter 30 Import substitution". *Handbook of Development Economics*. Ed. by Hollis Chenery {and} T. N. Srinivasan. Vol. 2. Elsevier, pp. 1601–1644.

Bruton, Henry J. (June 1, 1998). "A Reconsideration of Import Substitution". *Journal of Economic Literature* 36.2, pp. 903–936.

Bun, Maurice J. G. and Franc J. G. M. Klaassen (Apr. 3, 2002). *The Importance of Dynamics in Panel Gravity Models of Trade*. SSRN Scholarly Paper ID 306100. Rochester, NY: Social Science Research Network.

Carrothers, Gerald A. P. (June 30, 1956). "An Historical Review of the Gravity and Potential Concepts of Human Interaction". *Journal of the American Institute of Planners* 22.2, pp. 94–102.

Chamberlin, E. H. (1950). "Product Heterogeneity and Public Policy". *The American Economic Review* 40.2, pp. 85–92.

Chaney, Thomas (Aug. 2008). "Distorted Gravity: The Intensive and Extensive Margins of International Trade". *American Economic Review* 98.4, pp. 1707–1721.

Chenery, Hollis B., Sherman Robinson, and Moshe Syrquin (Oct. 1, 1986). *Industrialization and Growth : A Comparative Study*. The World Bank.

Chenery, Hollis B. and Moshe Syrquin (1986). "Typical Patterns of Transformation". *Industrialization and Growth : A Comparative Study*. Oxford University Press, pp. 37–83.

Chipman, John S. (1965). "A Survey of the Theory of International Trade: Part 1, The Classical Theory". *Econometrica* 33.3, pp. 477–519.

Clark, Don P. and W. Charles Sawyer (Sept. 2, 2014). "Stages of Diversification in Latin America". *Applied Economics Letters* 21.13, pp. 893–897.

Clark, Ximena, David Dollar, and Alejandro Micco (Dec. 2004). "Port Efficiency, Maritime Transport Costs, and Bilateral Trade". *Journal of Development Economics* 75.2, pp. 417–450.

Clark, Ximena, Timothy J Hatton, and Jeffrey G Williamson (Apr. 19, 2007). “Explaining U.S. Immigration, 1971–1998”. *Review of Economics and Statistics* 89.2, pp. 359–373.

Collier, Paul and Anke Hoeffler (2002). *Military Expenditure: Threats, Aid and Arms Races*. Policy, Research Working Paper WPS2927. Washington D.C: World Bank.

Condon, Timothy, Henrik Dahl, and Shantayanan Devarajan (1987). “Implementing a Computable General Equilibrium Model”. *World Bank DRC Papers* 290.

Davis, Graham A (Oct. 1, 1995). “Learning to Love the Dutch Disease: Evidence from the Mineral Economies”. *World Development* 23.10, pp. 1765–1779.

De Hoogh, Jerrie (1994). “Hans Linnemann: An Inspired and Inspiring Scientist”. *Trade, Aid and Development: Essays in Honour of Hans Linnemann*. Ed. by Jan Willem Gunning, Henk Kox, Wouter Tims, and Ynto de Wit. London: Palgrave Macmillan UK, pp. 1–14.

De Long, J. Bradford and Lawrence H. Summers (May 1, 1991). “Equipment Investment and Economic Growth”. *The Quarterly Journal of Economics* 106.2, pp. 445–502.

Deger, Saadet and Ron Smith (1983). “Military Expenditure and Growth in Less Developed Countries”. *The Journal of Conflict Resolution* 27.2, pp. 335–353.

Dennett, Adam (2016). “Estimating an Annual Time Series of Global Migration Flows - an Alternative Methodology for Using Migrant Stock Data”. *Global Dynamics: Approaches from Complexity Science*. Ed. by Alan Wilson. John Wiley and Sons, Chichester, pp. 125–141.

Dervis, Kemal, Jaime de Melo, and Sherman Robinson (Aug. 31, 1989). *General Equilibrium Models for Development Policy*. 10173. The World Bank.

Dhaene, Geert and Anton P. Barten (1989). “When It All Began: The 1936 Tinbergen Model Revisited”. *Economic Modelling* 6.2, pp. 203–219.

Dietzenbacher, Erik, Bart Los, Robert Stehrer, Marcel Timmer, and Gaaitzen de Vries (2013). “The Construction of World Input-Output Tables in the Wiod Project”. *Economic Systems Research* 25.1, pp. 71–98.

Disdier, Anne-Célia and Keith Head (2008). “The Puzzling Persistence of the Distance Effect on Bilateral Trade”. *The Review of Economics and Statistics* 90.1, pp. 37–48.

Dixit, Avinash K. and Joseph E. Stiglitz (1977). “Monopolistic Competition and Optimum Product Diversity”. *The American Economic Review* 67.3, pp. 297–308.

Dixon, Peter B. and B. R. Parmenter (Jan. 1, 1996). “Computable General Equilibrium Modelling for Policy Analysis and Forecasting”. *Handbook of Computational Economics* 1, pp. 3–85.

Dixon, Peter B. and Maureen T. Rimmer (May 2016). "Johansen's Legacy to CGE Modelling: Originator and Guiding Light for 50 Years". *Journal of Policy Modeling* 38.3, pp. 421–435.

Dornbusch, R. (1975). "Exchange Rates and Fiscal Policy in a Popular Model of International Trade". *The American Economic Review* 65.5, pp. 859–871.

Dornbusch, R., S. Fischer, and P. A. Samuelson (1977). "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods". *The American Economic Review* 67.5, pp. 823–839.

Dosi, Giovanni, Marco Grazzi, Luigi Marengo, and Simona Settepanella (2014). "Production theory: accounting for firm heterogeneity and technical change".

Duchin, Faye (2004). "International trade: evolution in the thought and analysis of Wassily Leontief". *Wassily Leontief and Input-Output Economics*, pp. 47–64.

Dudley, Leonard and Claude Montmarquette (1976). "A Model of the Supply of Bilateral Foreign Aid". *The American Economic Review* 66.1, pp. 132–142.

Dunne, J. Paul, Ron P. Smith, and Dirk Willenbockel (Jan. 1, 2005). "Models of Military Expenditure and Growth: A Critical Review". *Defence and Peace Economics* 16.6, pp. 449–461.

Durbarry, Ramesh (Dec. 1, 2004). "Tourism and economic growth: the case of Mauritius". *Tourism Economics* 10.4, pp. 389–401.

Eckstein, Susan (2001). *Power and Popular Protest: Latin American Social Movements*. University of California Press. 448 pp.

Economist, The (Mar. 10, 2012). "The end of cheap China". *The Economist*.

Edwards, Sebastian (July 1, 1992). "Trade Orientation, Distortions and Growth in Developing Countries". *Journal of Development Economics* 39.1, pp. 31–57.

Egger, Peter H., Maximilian von Ehrlich, and Douglas R. Nelson (Feb. 1, 2012). "Migration and Trade". *The World Economy* 35.2, pp. 216–241.

Egger, Peter and Michael Pfaffermayr (Mar. 1, 2004). "Distance, trade and FDI: a Hausman–Taylor SUR approach". *Journal of Applied Econometrics* 19.2, pp. 227–246.

Eifert, Benn, Alan Gelb, and Vijaya Ramachandran (Sept. 1, 2008). "The Cost of Doing Business in Africa: Evidence from Enterprise Survey Data". *World Development* 36.9, pp. 1531–1546.

eurostat (2015). *Reference And Management Of Nomenclatures*. URL: `http://ec.europa.eu/eurostat/ramon` (visited on 11/09/2015).

Fagan, Gabriel, Jérôme Henry, and Ricardo Mestre (Jan. 1, 2005). "An Area-Wide Model for the Euro Area". *Economic Modelling* 22.1, pp. 39–59.

Fagiolo, Giorgio and Marina Mastrorillo (2013). "Migration and Trade: A Complex-Network Approach". *arXiv preprint arXiv:1309.5859*.

Feder, Gershon (1983). "On Exports and Economic Growth". *Journal of development economics* 12 (1-2), pp. 59–73.

Feenstra, Robert C. (Nov. 10, 2015). *Advanced International Trade: Theory and Evidence, Second Edition*. Princeton University Press. 492 pp.

Feenstra, Robert C., Robert E. Lipsey, Haiyan Deng, Alyson C. Ma, and Hengyong Mo (2005). *World trade flows: 1962-2000*. National Bureau of Economic Research.

Felbermayr, Gabriel J. and Benjamin Jung (Aug. 2009). "The pro-trade effect of the brain drain: Sorting out confounding factors". *Economics Letters* 104.2, pp. 72–75.

Felbermayr, Gabriel J. and Farid Toubal (May 2012). "Revisiting the Trade-Migration Nexus: Evidence from New OECD Data". *World Development* 40.5, pp. 928–937.

Francois, Joseph F., Hans van Meijl, and Frank W. van Tongeren (2003). *Economic Benefits of the Doha Round for the Netherlands*. The Hague: Agricultural Economics Research Institute (LEI).

Francois, Joseph and Bernard Hoekman (Sept. 2010). "Services Trade and Policy". *Journal of Economic Literature* 48.3, pp. 642–692.

Francois, Joseph and Miriam Manchin (June 2013). "Institutions, Infrastructure, and Trade". *World Development* 46, pp. 165–175.

Frankel, Jeffrey A. (1997). *Regional Trading Blocs in the World Economic System*. Washington D.C: Institute for International Economics.

Fu, Xiaowen, Tae Hoon Oum, and Anming Zhang (2010). "Air Transport Liberalization and Its Impacts on Airline Competition and Air Passenger Traffic". *Transportation Journal* 49.4, pp. 24–41.

Gallup, John Luke, Jeffrey D. Sachs, and Andrew D. Mellinger (Aug. 1, 1999). "Geography and Economic Development". *International Regional Science Review* 22.2, pp. 179–232.

Gardiner, Ben, Ron Martin, Peter Sunley, and Peter Tyler (2013). "Spatially Unbalanced Growth in the British Economy". *Journal of Economic Geography* 13.6, pp. 889–928.

Genc, Murat, Masood Gheasi, Peter Nijkamp, and Jacques Poot (2012). "The impact of immigration on international trade: a meta-analysis". *Migration Impact Assessment: New Horizons*, p. 301.

Gibbs, Jack P. and Dudley L. Poston (1975). "The Division of Labor: Conceptualization and Related Measures". *Social Forces* 53.3, pp. 468–476.

Gibler, Douglas M. (2009). *International Military Alliances, 1648-2008*. CQ Press.

Girma, Sourafel and Zhihao Yu (Mar. 2002). "The link between immigration and trade: Evidence from the United Kingdom". *Weltwirtschaftliches Archiv* 138.1, pp. 115–130.

Gleditsch, Kristian Skrede (Jan. 10, 2002). “Expanded Trade and GDP Data”. *Journal of Conflict Resolution* 46.5, pp. 712–724.

Goldstein, Morris and Mohsin S. Khan (Jan. 1, 1985). “Chapter 20 Income and Price Effects in Foreign Trade”. *Handbook of International Economics* 2, pp. 1041–1105.

Goodwin, Richard M. (1949). “The Business Cycle as a Self-Sustaining Oscillation”. *Econometrica* 17.1949, pp. 184–185.

Gould, David M. (May 1994). “Immigrant Links to the Home Country: Empirical Implications for U.S. Bilateral Trade Flows”. *The Review of Economics and Statistics* 76.2, pp. 302–316.

Grogger, Jeffrey and Gordon H. Hanson (May 2011). “Income maximization and the selection and sorting of international migrants”. *Journal of Development Economics*. Symposium on Globalization and Brain Drain 95.1, pp. 42–57.

Grossman, Gene M. and Elhanan Helpman (1995). “Technology and Trade”. *Handbook of international economics* 3, pp. 1279–1337.

Hahsler, Michael, Christian Buchta, Kurt Hornik, Fionn Murtagh, Michael Brusco, Stephanie Stahl, and Hans-Friedrich Koehn (Aug. 23, 2015). *seriation: Infrastructure for Seriation*. Version 1.1-2.

Hanoch, Giora (1975). “Production and Demand Models with Direct or Indirect Implicit Additivity”. *Econometrica* 43.3, pp. 395–419.

Harrigan, Frank J., James W. McGilvray, and Ian H. McNicoll (1980). “Simulating the Structure of a Regional Economy”. *Environment and Planning A* 12.8, pp. 927–936.

Hausman, Warren, Hau Lee, and Uma Subramanian (2005). *Global Logistics Indicators, Supply Chain Metrics, and Bilateral Trade Patterns*. Policy Research Working Paper 3773. World Bank.

Helble, Matthias, Catherine L. Mann, and John S. Wilson (June 2012). “Aid-for-trade facilitation”. *Review of World Economics* 148.2, pp. 357–376.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein (May 1, 2008). “Estimating Trade Flows: Trading Partners and Trading Volumes”. *The Quarterly Journal of Economics* 123.2, pp. 441–487.

Heo, Uk (1998). “Modeling the Defense-Growth Relationship around the Globe”. *The Journal of Conflict Resolution* 42.5, pp. 637–657.

Hertel, Thomas W. (1997). *Global Trade Analysis: Modeling and Applications*. Cambridge University Press. 428 pp.

Hertel, Thomas W., J. M. Horridge, and K. R. Pearson (Oct. 1, 1992). “Mending the Family Tree: A Reconciliation of the Linearization and Levels Schools of AGE Modelling”. *Economic Modelling* 9.4, pp. 385–407.

Hertel, Thomas W. and Marinos E. Tsigas (1996). “Structure of GTAP”. *Global Trade Analysis*. Ed. by Thomas W. Hertel. Cambridge: Cambridge University Press, pp. 13–73.

Hertel, Thomas W., Terrie Walmsley, and Ken Itakura (2001). “Dynamic Effects of the" New Age" Free Trade Agreement between Japan and Singapore”. *Journal of Economic Integration*, pp. 446–484.

Hickman, Bert G. (1973). “A General Linear Model of World Trade”. *The International Linkage of National Economic Models*. Ed. by R. J. Ball. North-Holland, pp. 21–43.

Hickman, Bert G. and Lawrence J. Lau (1973). “Elasticities of Substitution and Export Demands in a World Trade Model”. *European Economic Review* 4.4, pp. 347–380.

Hickman, Bert G. and Stefan Schleicher (1978). “The Interdependence of National Economies and the Synchronization of Economic Fluctuations: Evidence from the LINK Project”. *Weltwirtschaftliches Archiv* 114.4, pp. 642–708.

Hicks, J. R. (1949). “Mr. Harrod’s Dynamic Theory”. *Economica* 16.62, pp. 106–121.

— (1950). *A Contribution to the Theory of the Trade Cycle*. At The Clarendon Press; Oxford.

Hidalgo, César A. and Ricardo Hausmann (2009). “The Building Blocks of Economic Complexity”. *Proceedings of the National Academy of Sciences* 106.26, pp. 10570–10575.

Hirschman, Albert O (1964). “The Paternity of an Index”. *The American Economic Review*, pp. 761–762.

Hirschman, Albert O. (1958). *The Strategy of Economic Development*. Yale University Press. 242 pp.

Hoekman, Bernard and Alessandro Nicita (Dec. 2011). “Trade Policy, Trade Costs, and Developing Country Trade”. *World Development* 39.12, pp. 2069–2079.

Hoover, Kevin D. and Stephen J. Perez (1999). “Data Mining Reconsidered: Encompassing and the General-to-Specific Approach to Specification Search”. *The Econometrics Journal* 2.2, pp. 167–191.

Houthakker, H. S. and Stephen P. Magee (1969). “Income and Price Elasticities in World Trade”. *The Review of Economics and Statistics* 51.2, pp. 111–125.

Howe, Howard (July 1, 1975). “Development of the Extended Linear Expenditure System from Simple Saving Assumptions”. *European Economic Review* 6.3, pp. 305–310.

Hummels, David and Georg Schaur (2012). *Time as a Trade Barrier*. National Bureau of Economic Research.

Imbs, Jean and Romain Wacziarg (Mar. 2003). “Stages of Diversification”. *American Economic Review* 93.1, pp. 63–86.

International Monetary Fund, ed. (2009). *Balance of payments and international investment position manual*. 6th ed. Washington D.C: International Monetary Fund. 351 pp.

Italianer, A. (1986). *Theory and Practice of International Trade Linkage Models*. Martin Nijhoff Publishers. 406 pp.

Iwanow, Tomasz and Colin Kirkpatrick (June 2009). "Trade Facilitation and Manufactured Exports: Is Africa Different?" *World Development* 37.6, pp. 1039–1050.

Jayaraman, T. K. and O. L. Shrestha (Dec. 1, 1976). "Some Trade Problems of Landlocked Nepal". *Asian Survey* 16.12, pp. 1113–1123.

Johansen, Leif (1960). *A Multi-Sectoral Study of Economic Growth*. North-Holland Publishing Company. 208 pp.

Johansen, Leif, Håvard Alstadheim, and Asmund Langsether (1968). "Explorations in Long-Term Projections for the Norwegian Economy". *Economics of Planning* 8.1, pp. 70–117.

Jones, Charles I. (Jan. 2011). *Misallocation, Economic Growth, and Input-Output Economics*. Working Paper 16742. National Bureau of Economic Research.

Karamba, Wendy R., Esteban J. Quiñones, and Paul Winters (Feb. 2011). "Migration and food consumption patterns in Ghana". *Food Policy*. Assessing the Impact of Migration on Food and Nutrition Security 36.1, pp. 41–53.

Keeney, Roman, Badri Narayanan, and Ernesto Valenzuela (2017). "The Global Trade Analysis Project's Database and CGE Model as a Tool for Agricultural and Environmental Economic Analysis". *The WSPC Reference on Natural Resources and Environmental Policy in the Era of Global Change*. World Scientific, pp. 31–56.

Kildegaard, Arne and Pete Williams (2002). "Banks, Systematic Risk, and Industrial Concentration: Theory and Evidence". *Journal of economic behavior & organization* 47.4, pp. 345–358.

Kim, Keuntae and Joel E. Cohen (Dec. 2010). "Determinants of International Migration Flows to and from Industrialized Countries: A Panel Data Approach Beyond Gravity1: Determinants of International Migration Flows". *International Migration Review* 44.4, pp. 899–932.

Kimura, Fukunari and Hyun-Hoon Lee (2006). "The Gravity Equation in International Trade in Services". *Review of World Economics / Weltwirtschaftliches Archiv* 142.1, pp. 92–121.

Kindleberger, C. P. (1952). "Reviews of Papers on Turkey, Guatemala and Cuba". *The Review of Economics and Statistics* 34.4, pp. 391–394.

Klein, Lawrence R. and Herman Rubin (1947). "A Constant-Utility Index of the Cost of Living". *The Review of Economic Studies* 15.2, pp. 84–87.

Klein, Lawrence R. and A. Van Peeterssen (1973). “Forecasting World Trade within Project LINK”. *The International Linkage of National Economic Models*. Ed. by R. J. Ball. North-Holland, pp. 429–463.

Koopman, Robert, Zhi Wang, and Shang-Jin Wei (2008). *How much of Chinese exports is really made in China? Assessing domestic value-added when processing trade is pervasive*. National Bureau of Economic Research.

Korinek, Jane and Patricia Sourdin (2009). “Maritime Transport Costs and Their Impact on Trade”. *OECD* Report code TAD/TC/WP (2009)7.

Krugman, Paul (1980a). “Scale Economies, Product Differentiation, and the Pattern of Trade”. *The American Economic Review* 70.5, pp. 950–959.

— (Dec. 1, 1980b). “Scale Economies, Product Differentiation, and the Pattern of Trade”. *The American Economic Review* 70.5, pp. 950–959.

— (1989). “Differences in Income Elasticities and Trends in Real Exchange Rates”. *European Economic Review* 33.5, pp. 1031–1046.

Kubo, Yuji, Jaime De Melo, Sherman Robinson, and Moshe Syrquin (1986). “Interdependence and Industrial Structure”. *Industrialization and Growth : A Comparative Study*. Oxford University Press, pp. 188–225.

Lahr, Michael and Louis de Mesnard (June 2004). “Biproportional Techniques in Input-Output Analysis: Table Updating and Structural Analysis”. *Economic Systems Research* 16.2, pp. 115–134.

League of Nations (1942). *The Network of World Trade: A Companion Volume to "Europe’s Trade."* Geneva: League of Nations.

Lenzen, Manfred, Keiichiro Kanemoto, Daniel Moran, and Arne Geschke (Aug. 7, 2012). “Mapping the Structure of the World Economy”. *Environmental Science & Technology* 46.15, pp. 8374–8381.

Lenzen, Manfred, Daniel Moran, Keiichiro Kanemoto, and Arne Geschke (2013). “Building Eora: A Global Multi-Region Input-Output Database at High Country and Sector Resolution”. *Economic Systems Research* 25.1, pp. 20–49.

Leontief, Wassily (1949). “Recent Developments in the Study of Interindustrial Relationships”. *The American Economic Review* 39.3, pp. 211–225.

Leontief, Wassily W. (Aug. 1, 1936). “Quantitative Input and Output Relations in the Economic Systems of the United States”. *The Review of Economics and Statistics* 18.3, pp. 105–125.

Levy, Robert G., Thomas P. Oléron Evans, and Alan Wilson (2016). “A Global Inter-Country Economic Model Based on Linked Input–Output Models”. *Global Dynamics*. Ed. by Alan Wilson. Wiley, pp. 51–72.

Lewis, William Arthur (1955). *The Theory of Economic Growth*. R.D. Irwin. 464 pp.

Li, Hongbin, Lei Li, Binzhen Wu, and Yanyan Xiong (Nov. 2012). "The End of Cheap Chinese Labor". *Journal of Economic Perspectives* 26.4, pp. 57–74.

Limão, Nuno and Anthony J. Venables (2001). "Infrastructure, Geographical Disadvantage, Transport Costs, and Trade". *The World Bank Economic Review* 15.3, pp. 451–479.

Linders, Gert-Jan M., Arjen HL Slangen, Henri L.F. De Groot, and Sjoerd Beugelsdijk (2005). *Cultural and Institutional Determinants of Bilateral Trade Flows*. Tinbergen Institute Discussion Paper.

Linnemann, Hans (1966). *An Econometric Study of International Trade Flows*. Vol. 234. North-Holland Publishing Company Amsterdam.

Lipton, Michael (1962). "Balanced and Unbalanced Growth in Underdeveloped Countries". *The Economic Journal* 72.287, pp. 641–657.

Lucas, Robert E. (1988). "On the Mechanics of Economic Development". *Journal of Monetary Economics* 22, pp. 3–42.

Maisel, Sebastian and John A. Shoup (Feb. 2009). *Saudi Arabia and the Gulf Arab States Today: An Encyclopedia of Life in the Arab States*. Greenwood Publishing Group. 606 pp.

Mankiw, N. Gregory, David Romer, and David N. Weil (1992). "A Contribution to the Empirics of Economic Growth". *The Quarterly Journal of Economics* 107.2, pp. 407–437.

Markusen, James R. (July 2013). "Putting per-capita income back into trade theory". *Journal of International Economics* 90.2, pp. 255–265.

Marquez, Jaime (1990). "Bilateral Trade Elasticities". *The Review of Economics and Statistics* 72.1, pp. 70–77.

Martínez-Zarzoso, Inmaculada (Jan. 1, 2013). "The Log of Gravity Revisited". *Applied Economics* 45.3, pp. 311–327.

Marwah, Kanta (1976). "A World Model of International Trade: Forecasting Market Shares and Trade Flows". *Empirical Economics* 1.1, pp. 1–39.

Marwah, Kanta and Romesh Diwan (1975). "Structural Shifts Subsidies and Non-Market Factors in International Trade". *Eastern Economic Journal* 2.3, pp. 212–234.

Mayda, Anna Maria (May 28, 2009). "International migration: a panel data analysis of the determinants of bilateral flows". *Journal of Population Economics* 23.4, pp. 1249–1274.

McCallum, John (June 1, 1995). "National Borders Matter: Canada-U.S. Regional Trade Patterns". *The American Economic Review* 85.3, pp. 615–623.

McDougall, Robert (2003). "A New Regional Household Demand System for GTAP". GTAP Technical Papers.

McGillivray, Mark and Edward Oczkowski (1992). “A Two-Part Sample Selection Model of British Bilateral Foreign Aid Allocation”. *Applied Economics* 24.12, pp. 1311–1319.

McKinlay, Robert D. and Richard Little (1977). “A Foreign Policy Model of US Bilateral Aid Allocation”. *World Politics* 30.1, pp. 58–86.

Melitz, Jacques (May 2008). “Language and foreign trade”. *European Economic Review* 52.4, pp. 667–699.

Melitz, Marc J. (Nov. 1, 2003). “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity”. *Econometrica* 71.6, pp. 1695–1725.

Metzler, Lloyd A. (1950). “A Multiple-Region Theory of Income and Trade”. *Econometrica* 18.4, pp. 329–354.

Miller, Marcus H. and John E. Spencer (Feb. 1977). “The Static Economic Effects of the UK Joining the EEC: A General Equilibrium Approach”. *The Review of Economic Studies* 44.1, p. 71.

Miller, Ronald E. and Peter D. Blair (1985). *Input-output analysis: foundations and extensions*. Prentice-Hall. 488 pp.

— (July 30, 2009). *Input-Output Analysis: Foundations and Extensions*. Cambridge University Press. 783 pp.

Minor, Peter and Marinos Tsigas (2008). “Impacts of Better Trade Facilitation in Developing Countries”. Submitted to the GTAP 11th Annual Conference. Helsinki, Finland.

Mosteller, Frederick (Mar. 1, 1968). “Association and Estimation in Contingency Tables”. *Journal of the American Statistical Association* 63.321, pp. 1–28.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny (1989). “Industrialization and the Big Push”. *Journal of Political Economy* 97.5, pp. 1003–1026.

Neisser, Hans and Franco Modigliani (1953). “National Incomes and International Trade”.

Newman, M. E. J. and M. Girvan (Feb. 26, 2004a). “Finding and evaluating community structure in networks”. *Physical Review E* 69.2, p. 026113.

Newman, Mark EJ (2010). *Networks: An Introduction*. Oxford University Press.

Newman, Mark EJ and Michelle Girvan (2004b). “Finding and Evaluating Community Structure in Networks”. *Physical review E* 69.2, p. 026113.

Nordhaus, William, John R. Oneal, and Bruce Russett (2012). “The Effects of the International Security Environment on National Military Expenditures: A Multicountry Study”. *International Organization* 66.3, pp. 491–513.

Novy, Dennis (2013). “Gravity Redux: Measuring International Trade Costs with Panel Data”. *Economic Inquiry* 51.1, pp. 101–121.

Nurske, Ragnar (1953). *Problems of Capital Formation In Underdeveloped Countries*. 2nd ed. Oxford: Basil Blackwell.

Nystuen, John D. and Michael F. Dacey (1961). "A Graph Theory Interpretation of Nodal Regions". *Papers of the Regional Science Association*. Vol. 7, pp. 29–42.

ONS (Aug. 23, 2013). *Understanding GDP and how it is measured*. URL: `http://www.ons.gov.uk/ons/rel/elmr/explaining-economic-statistics/understanding-gdp-and-how-it-is-measured/sty-understanding-gdp.html` (visited on 07/17/2015).

Oosterhaven, Jan (Aug. 1, 1989). "The Supply-Driven Input-Output Model: A New Interpretation but Still Implausible". *Journal of Regional Science* 29.3, pp. 459–465.

— (Mar. 1, 2012). "Adding Supply-Driven Consumption Makes the Ghosh Model Even More Implausible". *Economic Systems Research* 24.1, pp. 101–111.

Ortega, Francesc and Giovanni Peri (Jan. 3, 2013). "The effect of income and immigration policies on international migration". *Migration Studies* 1.1, pp. 47–74.

Otsuki, Tsunehiro, John S. Wilson, and Mirvat Sewadeh (2001). "Saving Two in a Billion:: Quantifying the Trade Effect of European Food Safety Standards on African Exports". *Food policy* 26.5, pp. 495–514.

Panne, C. and G. J. Averink (1961). "Imperfect Management Decisions and Predictions and Their Financial Implications in Dynamic Quadratic Cost Minimization". *Statistica Neerlandica* 15.3, pp. 293–318.

Parikh, Ashok (1979). "Forecasts of Input-Output Matrices Using the R.A.S. Method". *The Review of Economics and Statistics* 61.3, pp. 477–481.

Pearson, K. R. (1988). "Automating the Computation of Solutions of Large Economic Models". *Economic Modelling* 5.4, pp. 385–395.

Pedersen, Peder J., Mariola Pytlikova, and Nina Smith (Oct. 2008). "Selection and network effects—Migration flows into OECD countries 1990–2000". *European Economic Review* 52.7, pp. 1160–1186.

Polak, J. J. (1939). "International Propagation of Business Cycles". *The Review of Economic Studies* 6.2, pp. 79–99.

— (1954). *An International Economic System*. Chicago University Press. 181 pp.

Polenske, Karen R. (1997). "Current Uses of the RAS Technique: A Critical Review". *Prices, Growth and Cycles*. Palgrave Macmillan, London, pp. 58–88.

Pollak, Robert A. and Terence J. Wales (1969). "Estimation of the Linear Expenditure System". *Econometrica* 37.4, pp. 611–628.

Portugal-Perez, Alberto and John S. Wilson (July 2012). "Export Performance and Trade Facilitation Reform: Hard and Soft Infrastructure". *World Development* 40.7, pp. 1295–1307.

Pöyhönen, Pentti (1963). "A Tentative Model for the Volume of Trade between Countries". *Weltwirtschaftliches Archiv* 90, pp. 93–100.

Prais, S. J. (Aug. 1, 1962). "Econometric Research in International Trade: A Review". *Kyklos* 15.3, pp. 560–574.

Pursell, Lyle and S. Y. Trimble (1991). "Gram-Schmidt Orthogonalization by Gauss Elimination". *The American Mathematical Monthly* 98.6, pp. 544–549.

Radelet, Steven and Jeffrey Sachs (1998). "Shipping Costs, Manufactured Exports, and Economic Growth". *American Economic Association Meetings, Harvard University, Mimeo*.

Ram, Rati (1986). "Government Size and Economic Growth: A New Framework and Some Evidence from Cross-Section and Time-Series Data". *The American Economic Review* 76.1, pp. 191–203.

— (1995). "Defense Expenditure and Economic Growth". *Handbook of Defense Economics* 1, pp. 251–274.

Ramalingam, Ben, Harry Jones, Toussaint Reba, and John Young (Oct. 20, 2009). *Exploring the science of complexity: Ideas and implications for development and humanitarian efforts*. URL: `http://dspace.cigilibrary.org/jspui/handle/123456789/25084` (visited on 11/19/2012).

Raymer, James, Arkadiusz Wiśniowski, Jonathan J. Forster, and Peter W. F. Smith (Sept. 2013). "Integrated Modeling of European Migration". *Journal of the American Statistical Association* 108.503, pp. 801–819.

Reed, Wallace E. (1967). *Areal Interaction in India: Commodity Flows of the Bengal-Bihar Industrial Area*. 110. University of Chicago, Committee on Geographical Studies, Research Papers.

Reichardt, Jörg and Stefan Bornholdt (2006a). "Statistical Mechanics of Community Detection". *Physical Review E* 74.1, p. 016110.

— (July 18, 2006b). "Statistical mechanics of community detection". *Physical Review E* 74.1, p. 016110.

Reilly, William John (1931). *The Law of Retail Gravitation*. W.J. Reilly. 94 pp.

Rhomberg, Rudolf R. (1966a). "A Short-Term World Trade Model". *Econometrica* 34.5, pp. 90–91.

— (1966b). "Review of Measuring Transactions between World Areas". *The American Economic Review* 56.5. In collab. with Herbert B. Woolley, pp. 1325–1327.

Rhomberg, Rudolf R. (1970). "Possible Approaches to a Model of World Trade and Payments (Méthodes Possibles de Construction d'un Modèle Du Commerce et Des Paiements Mondiaux) (Posibles Enfoques de Un Modelo de Comercio y Pagos Mundiales)". *Staff Papers (International Monetary Fund)* 17.1, pp. 1–28.

— (1973). "Towards a General Trade Model". *The International Linkage of National Economic Models*. Ed. by R. J. Ball. North-Holland, pp. 9–20.

Richardson, Lewis Fry (1919). *Mathematical Psychology of War*. Oxford: William Hunt.

— (1960). *Arms and Insecurity*. Pittsburgh, PA: The Boxwood Press.

Robinson, Sherman (Jan. 1, 1989). "Chapter 18 Multisectoral Models". *Handbook of Development Economics*. Vol. 2. Elsevier, pp. 885–947.

— (2006). "Macro Models and Multipliers: Leontief, Stone, Keynes, and CGE Models". *Poverty, Inequality and Development*. Springer, pp. 205–232.

Rodrik, Dani (2015). *Economics Rules: The Rights and Wrongs of The Dismal Science*. New York: W.W. Norton.

Romer, Paul M. (1986). "Increasing Returns and Long-Run Growth". *Journal of political economy* 94.5, pp. 1002–1037.

Rose, Adam and William Miernyk (1989). "Input–output Analysis: The First Fifty Years". *Economic Systems Research* 1.2, pp. 229–272.

Rose, Andrew K. (May 1, 1991). "The Role of Exchange Rates in a Popular Model of International Trade: Does the 'Marshall–Lerner' Condition Hold?" *Journal of International Economics* 30.3, pp. 301–316.

Rosenstein-Rodan, P. N. (1943). "Problems of Industrialisation of Eastern and South-Eastern Europe". *The Economic Journal* 53 (210/211), pp. 202–211.

Rosvall, Martin and Carl T. Bergstrom (Jan. 29, 2008). "Maps of random walks on complex networks reveal community structure". *Proceedings of the National Academy of Sciences* 105.4, pp. 1118–1123.

Rutherford, Thomas F. (2005). "GTAP6inGAMS: The Dataset and Static Model". *Mimeo, University of Colorado*.

Sawyer, John A. (Nov. 1967). "An Econometric Study of International Trade Flows". *The Canadian Journal of Economics and Political Science* 33.4, p. 633.

Scarf, Herbert et al. (1967). *On the Computation of Equilibrium Prices*. 232. Cowles Foundation for Research in Economics at Yale University.

Shaikh, Anwar (Feb. 1974). "Laws of Production and Laws of Algebra: The Humbug Production Function". *The Review of Economics and Statistics* 56.1, p. 115.

Shannon, Lyle W. (Sept. 1, 1959). "Albert O. Hirschman:. The Strategy of Economic Development". *The Annals of the American Academy of Political and Social Science* 325.1, pp. 125–126.

Shin, M. and M.D. Ward (1999). "Lost in Space: Political Geography and the Defense-Growth Trade-Off". *Journal of Conflict Resolution* 43, pp. 793–817.

Shoven, John B. and John Whalley (1984). "Applied General-Equilibrium Models of Taxation and International Trade: An Introduction and Survey". *Journal of Economic Literature* 22.3, pp. 1007–1051.

Sierksma, Gerard (Aug. 1979). "Non-negative matrices: The open Leontief model". *Linear Algebra and its Applications* 26, pp. 175–201.

Signorino, Curtis S. and Jeffrey M. Ritter (1999). "Tau-b or Not Tau-b: Measuring the Similarity of Foreign Policy Positions". *International Studies Quarterly* 43.1, pp. 115–144.

Silva, JMC Santos and Silvana Tenreyro (2006). "The Log of Gravity". *The Review of Economics and statistics* 88.4, pp. 641–658.

Simoes, Alexander James Gaspar and César A. Hidalgo (2011). "The Economic Complexity Observatory: An Analytical Tool for Understanding the Dynamics of Economic Development." *Scalable Integration of Analytics and Visualization*.

Simpson, Edward H (1949). "Measurement of Diversity". *Nature* 163, pp. 688–688.

Smith, Robert H. T. (1970). "Concepts and Methods in Commodity Flow Analysis". *Economic Geography* 46, pp. 404–416.

Soloaga, Isidro, John S. Wilson, and Alejandro Mejia (June 1, 2006). *Moving forward faster : trade facilitation reform and Mexican competitiveness*. WPS3953. The World Bank, pp. 1–31.

Solow, Robert M. (1955). "The Production Function and the Theory of Capital". *The Review of Economic Studies* 23.2, pp. 101–108.

Stiglitz, Joseph (Sept. 11, 2007). *The Malaysian Miracle*. Project Syndicate. URL: `project-syndicate.org/commentary/the-malaysian-miracle` (visited on 08/14/2015).

Stone, Richard (1954). "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand". *The Economic Journal* 64.255, pp. 511–527.

Stone, Richard and Alan Brown (1962). *A Computable Model of Economic Growth*. Vol. 1. London: Chapman and Hall.

Streeten, Paul (1959). "Unbalanced Growth". *Oxford Economic Papers* 11.2, pp. 167–190.

Streicher, Gerhard and Robert Stehrer (2012). "Whither Panama? Constructing a consistent and balanced world SUT system including international trade and transport margins". *Institutions*.

Subramanian, Uma and John Arnold (Jan. 1, 2001). *Forging Subregional Links in Transportation and Logistics in South Asia*. World Bank Publications.

Suwa-Eisenmann, Akiko and Thierry Verdier (Sept. 21, 2007). "Aid and trade". *Oxford Review of Economic Policy* 23.3, pp. 481–507.

Taplin, Grant B. (1967). "Models of World Trade (Modèles de Commerce Mondial) (Modelos Sobre Comercio Mundial)". *Staff Papers (International Monetary Fund)* 14.3, pp. 433–455.

Temple, Jonathan (2008). "Balanced Growth". *The New Palgrave Dictionary of Economics*. Ed. by Lawrence E. Bloom and Steven N. Durlauf. London: Palgrave Macmillan UK, pp. 1–4.

The World Input-Output Database (WIOD): Contents, Sources and Methods (Apr. 2012). 10. World Input-Output Database.

Theil, Henri (1961). "The Econometric Institute's First Five Years". *De Economist* 109.9, pp. 603–634.

Thompson, Earl A. (July 1, 1974). "Taxation and National Defense". *Journal of Political Economy* 82.4, pp. 755–782.

Timmer, Marcel P., Erik Dietzenbacher, Bart Los, Robert Stehrer, and Gaaitzen J. de Vries (May 1, 2015). "An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production". *Review of International Economics*, n/a–n/a.

Timmer, Marcel, AA Erumban, R Gouma, B Los, U Temurshoev, GJ de Vries, and I Arto (2012). "The World Input-Output Database (WIOD): Contents, Sources and Methods".

Tinbergen, Jan (1936). *Kan Hier Te Lande, Al Dan Niet Na Overheidsingrijpen, Een Verbetering van de Binnenlandsche Conjunctuur Intreden, Ook Zonder Verbetering van Onze Exportpositie?* The Hague: Martinus Nijhoff.

— (1962). *Shaping the World Economy; Suggestions for an International Economic Policy*. New York: The Twentieth Century Fund.

Trumbull, William N. and Howard J. Wall (1994). "Estimating Aid-Allocation Criteria with Panel Data". *The Economic Journal*, pp. 876–882.

Tukker, Arnold and Erik Dietzenbacher (Mar. 2013). "Global Multiregional Input–Output Frameworks: an Introduction and Outlook". *Economic Systems Research* 25.1, pp. 1–19.

Tukker, Arnold, Arjan de Koning, Richard Wood, Troy Hawkins, Stephan Lutter, Jose Acosta, Jose M. Rueda Cantuche, Maaike Bouwmeester, Jan Oosterhaven, Thomas Drosdowski, and Jeroen Kuenen (2013). "Exiopol — Development and Illustrative Analyses of a Detailed Global MR EE SUT/IOT". *Economic Systems Research* 25.1, pp. 50–70.

UNDP (1992). *Human Development Report 1992*. Oxford University Press.

Venkateshan, SP and Prasanna Swaminathan (2013). *Computational Methods in Engineering*. Academic Press.

Waelbroeck, Jean (1964). “Le Commerce de La Communauté Européenne Avec Les Pays Tiers”. *Intégration Européene et Réalité Economique*.

— (1967). “On the Structure of International Trade Interdependence”. *Brussels Economic Review* 36, pp. 495–512.

Walkenhorst, Peter and Tadashi Yasui (2003). *Quantitative Assessment of the Benefits of Trade Facilitation*. Report number TD/TC/WP(2003)31. Paris: OECD.

Wallendorf, Melanie and Michael D. Reilly (Dec. 1, 1983). “Ethnic Migration, Assimilation, and Consumption”. *Journal of Consumer Research* 10.3, pp. 292–302.

Walmsley, Terrie, Angel Aguiar, and Badri Narayanan (2012). *Introduction to the Global Trade Analysis Project and the GTAP Data Base*. Working Paper 67. Center for Global Trade Analysis, Purdue University.

Wan, Michael (Jan. 22, 2013). *Electronic exports: Identifying Asia’s winners and losers*. Credit Suisse.

Ward, Michael D., John S. Ahlquist, and Arturas Rozenas (2013). “Gravity’s Rainbow: A dynamic latent space model for the world trade network”. *Network Science* 1.1, pp. 95–118.

Weidmann, Nils B., Doreen Kuse, and Kristian Skrede Gleditsch (Feb. 26, 2010). “The Geography of the International System: The CShapes Dataset”. *International Interactions* 36.1, pp. 86–106.

Wilson, A. G. (Jan. 1, 1969). “The Use of Entropy Maximising Models, in the Theory of Trip Distribution, Mode Split and Route Split”. *Journal of Transport Economics and Policy* 3.1, pp. 108–126.

Wilson, Alan (1971). “A family of spatial interaction models, and associated developments”. *Environment and Planning* 3.1, pp. 1–32.

Wilson, Alan G. (1967). “A statistical theory of spatial distribution models”. *Transportation research* 1.3, pp. 253–269.

Wilson, J. S., C. L. Mann, and T. Otsuki (Dec. 2, 2003). “Trade Facilitation and Economic Development: A New Approach to Quantifying the Impact”. *The World Bank Economic Review* 17.3, pp. 367–389.

Woolley, Herbert B. (1966). *Measuring Transactions between World Areas*. NBER Books. National Bureau of Economic Research, Inc.

Yakovlev, Pavel (Aug. 1, 2007). “Arms Trade, Military Spending, and Economic Growth”. *Defence and Peace Economics* 18.4, pp. 317–338.

Yang, Heewon, Chanyoung Hong, Sungmoon Jung, and Jeong-Dong Lee (July 1, 2015). “Arms or Butter: The Economic Effect of an Increase in Military Expenditure”. *Journal of Policy Modeling* 37.4, pp. 596–615.

Yeates, Maurice H. (Oct. 1, 1969). “A Note Concerning the Development of a Geographic Model of International Trade”. *Geographical Analysis* 1.4, pp. 399–404.

Yildirim, J. and N. Öcal (2014). “Military Expenditures, Economic Growth and Spatial Spillovers”. *Defence and Peace Economics* 87.1, pp. 87–104.

Younas, Javed (Sept. 1, 2008). “Motivation for Bilateral Aid Allocation: Altruism or Trade Benefits”. *European Journal of Political Economy* 24.3, pp. 661–674.

Young, E. C. (1924). *The Movement of Farm Population.* Cornell University. 91 pp.

Zaki, Chahir (2009). “Towards an Explicit Modeling of Trade Facilitation in CGE Models: Evidence from Egypt”. *Mimeo, Centre d’Economie de la Sorbonne, Paris*.