Georg-August-University of Göttingen

Doctoral Thesis

Essays on effects of policy interventions in the realm of food standards, trade, and the German labour market

---

Dissertation
zur Erlangung des Doktorgrades
der Wirtschaftswissenschaftlichen Fakultät
der Georg-August-Universität Göttingen

vorgelegt von
Malte Ehrich,
geboren in Mölln

Göttingen, 13. Februar 2017

Erstgutachter	Prof. Dr. Bernhard Brümmer
Zweitgutachterin:	Prof. Inmaculada Martínez-Zarzoso, PhD
Drittprüfer:	Prof. Dr. Thomas Kneib
Disputation:	24. März 2017

"*Give the ones you love wings to fly, roots to come back and reasons to stay.*"

Dalai Lama

"*You can always count on Americans to do the right thing - after they've tried everything else.*"

Winston Churchill

"*If economics were only about profit maximisation, it would be just another name for business administration. It is a social discipline, and society has other means of cost accounting besides market prices.*"

Dani Rodrik

Acknowledgements

Writing this thesis and finalising the PhD has been an amazing journey during which I have met great and unique characters. First of all, I would like to thank my supervisor Bernhard Brümmer. You continuously trust my work and skills more than I did myself. Based on your fundamental belief in intrinsic motivation of humans, you guarantee your students a degree of academic freedom which is exceptional. In case of mistakes or misfortune, you always defend your students such that I always felt being accepted and appreciated by you including all my strengths and weaknesses. I would also like to thank the other two members of the supervision committee Inmaculada Martínez-Zarzoso and Thomas Kneib for excellent supervision.

Furthermore, a range of additional researchers that have I met during this journey in various settings deserve deep thankfulness. Sebastian Hess, thank you for hosting me at my first academic research stay abroad in Uppsala. I am thankful for your academic and personal advises. If I had stayed in academia, you would have been responsible for a significant part of that. You have been a great mentor for me! Moreover, I am heavily indebted to my counter-part at IFPRI Devesh Roy. I will never forget our coffees and teas at Starbucks at Horniman circle in Mumbai during which we discussed and developed countless research ideas which resulted in our paper about the German labour market reforms. At a later stage, I was delighted to meet Anirudh Shingal from the World Trade Institute in Bern at the ETSG conference in Paris. The sometimes daily Skype-talks enriched my understanding on economics overall, but also on *how* to work in empirical trade analyses.

Another person deserves her individual paragraph. Melanie Grosse, thank you for your continuous and successful effort to improve the research environment for us (post-) doctoral researchers at GlobalFood every day.

The joy of teaching was one major motivation for me to start a PhD. Teaching and supervision of theses became a crucial ingredient of my PhD-life and I am deeply thankful to my students. You not only pulled me away from boring desk-work and endless Stata-coding, but more importantly our interactions and discussions gave me the feeling of doing something with significance. In particular, I would like to thank Dela, Nina, and Daniel for excellent work and highly valuable characters. I wish you all the best for your future!

Whenever I have been asked about my PhD-life in Göttingen, I have answered in the same way without any exception: meeting my colleagues and spending time with them at GlobalFood or privately belongs to the best experiences during these years. In particular I would like to thank my office mate Dirk. Unfortunately, I have never been capable enough to understand the deepness of your jokes. I will miss you and

them nevertheless. Moreover, Lutz and Kathrin next door, meeting you every day was a great pleasure and made my every-day life much more enjoyable. Finally, Eva, Kerstin, Andrea, Lisa, Marry, Katharina, Hanna, Marwan, Jana, Marion, Nathalie, Schneidi, Insa, and Hendrik: Thanks for accompanying me and for amazing friendships.

I have never considered the PhD-degree to be purely a result of hard work of an individual with a sufficient mindset. I grew up with a privileged background which was shaped by my family but also by society overall. As a child of two teachers, I have never experienced constraints due to income and I have always been aware of the value of good education. Thus, I am thankful to my mother Wiebke and my father Claus for providing me the suitable environment that allowed me to develop freely and independently a critical mindset. This is not every child's privilege in Germany. I am also thankful to my brother Thies and my sister Inken who have always been – without realising – ideals from early childhood onward.

How valuable a PhD-degree ever might be, it will never reach the significance of sharing my life with the most wonderful person I have ever met. Thank you Chrissy for your love and for making me realising every single day what remains important in life.

Contents

Overview of co-authors

- **Brümmer**, Prof. Dr. Bernhard: Professor of Agricultural Market Analysis, *Georg-August-University Göttingen, Germany*
  Chapter 3
- **Hess**, Prof. Dr. Sebastian: Professor of Dairy Economics
  Christian-Albrechts-University Kiel, Germany
  Chapter 4
- **Mangelsdorf**, Dr. Axel: Economist and Consultant
  GFA Group Hamburg, Technical University Berlin, The World Bank
  Chapter 5
- **Martínez-Zarzoso**, Prof. Inmaculada (PhD): Professor of Empirical Trade Analysis, *Georg-August-University Göttingen, Germany*
  Chapter 3
- **Munasib**, Abdul (Phd): Economist
  Bureau of Economic Analysis (BEA), USA
  Chapter 6
- **Roy**, Devesh (Phd): Senior Research Fellow
  International Food Policy Research Institute (IFPRI), New-Delhi
  Chapter 6

List of Figures

List of Tables

List of Abbreviations

ALG	**A**rbeitslosengeld (unemployment benefit)
APEMR	**A**bsolute **P**rediction **E**rror to **M**ean **R**atio
B2B	**B**usiness **t**o **B**usiness
B2C	**B**usiness **t**o **C**onsumers
BRC	**B**ritish **R**etail **C**onsortium
CEPII	**C**entre d'**E**tudes **P**rospectives et d'**I**nformations **I**nternationales
FAO	**F**ood and **A**gricultural **O**rganization
FLO	**F**ederal **L**abour **O**ffice
GATT	**G**eneral **A**greement on **T**ariffs and **Trade**
GDP	**G**ross **D**omestic **P**roduct
GMO	**G**enetically **M**odified **O**rganism
HACCP	**H**azard **A**nalysis **C**ritical **C**ontrol **P**oints
HS	**H**armonisation **S**ystem
IFS	**I**nternational **F**eatured **S**tandards
ITC	**I**nternational **T**rade **C**enter
LFP	**L**abour **F**orce **P**articipation
MSC	**M**arine **S**tewardship **C**ouncil
MRL	**M**aximum **R**esidue **L**imits
NTB	**N**on **T**ariff **B**arriers
OECD	**O**rganisation for **E**conomic **C**ooperation and **Development**
ppm	**p**arts **per**million
PCA	**P**rincipal **C**omponent**A**nalysis
PSA	**P**ersonal **S**ervice**A**gency
PPML	**P**oisson **P**seudo **M**aximum **Likelihood**
RMSPE	**R**oot **M**ean **S**quared **P**rediction **E**rror
SAA	**S**tandard **A**doption**A**bility
SCM	**S**ynthetic **C**ontrol**M**ethod
SPS	**S**anitary and **P**hyto**s**anitary
TBT	**T**echnical **B**arriers to **T**rade
TTIP	**T**ransatlantic **T**rade and **I**nvestment **P**artnership
WDI	**W**orld **D**evelopment **I**ndicators
WGI	**W**orldwide **G**overnance **I**ndicators
WHO	**W**orld **H**ealth **O**rganization
WITS	**W**orld **I**ntegrated **T**rade **S**olution
WTO	**W**orld **T**rade **O**rganization
WTP	**W**illingess **t**o **P**ay

Für Chrissy

Preface

International trade and the integration of global markets are important determinants of individual wealth and well-being. The main channels through which trade affects welfare are increasing variety of purchasable products, lower consumer prices, and higher income levels on average. The effects of trade on various outcomes such as income, welfare, or inequality depend massively on the specific context and institutional environment such that trade should not be analysed in isolation. Because of the highly relevant consequences of trade, the factors which determine the functionality of trade as a wealth-generating and re-distributing machinery are important aspects to analyse.

This doctoral thesis is written in the early stage of an era which is likely to be determined by raising popularity of anti-liberal and anti-trade advocates as impressively demonstrated by the recently elected president of the United States of America. As one of his first acts, the president of the US signed the withdrawal from the agreement to form the Trans-Pacific Partnership which was initiated as a broad regional trade agreement connecting the USA with many East-Asian states. Moreover, he suggested to introduce punitive tariffs for particular products aiming to protect domestic industries and jobs (Morici, 2017). Such policies do not only ignore but also antagonise economic laws. The results will probably be contrary to the policy objectives which are official stated such as higher economic growth rates and reduction of unemployment.[1] As a response, the European Union (EU) has begun to prepare a trade war in case protectionist policies of the US intensify (Schieritz, 2017).

The American president is not alone. The Turkish parliament has voted for its own abolishment by accepting a reform which increases power and authority of the president. The British have voted for leaving the EU - one of the most successful political projects (Rodrik, 2012) - in order to regain complete national sovereignty.[2] Marine Le Pen, the leader of the right-extreme Front National in France, could win French elections in April 2017. The most recent poll from 2^{nd} February 2017 forecasts the "Alternative für Deutschland" AfD (right-extreme party) to gain 12% if elections had taken place that day. A nationalist and anti-liberal party would rank third in a general German election. This is a novelty in German politics after World War II. This development goes hand-in-hand with increasing nationalism and higher voting shares of anti-liberal parties in other countries as well as it is the case for example in Poland, Denmark, Sweden, United Kingdom, and probably soon in France, Netherlands, and Germany. A doctoral thesis within the broad field of globalisation and international

[1] However, these objectives do not necessarily need to coincide with the *de facto* objectives of the government which remain speculative.

[2] Rodrik re-considered his view about the EU in subsequent years as response to insufficient EU-policies during the economic and institutional crises of the EU (Zajonz, 2016).

trade can not ignore the current raising popularity of protectionist trade policies although the explicit research questions themselves address these concerns only partly if at all. It is my concern to embed my research on trade in the recent political development.

If this trend of raising nationalism continuous or even intensifies, global trade and income of particular countries are likely to decline with significant distributional consequences. Unemployment is likely to increase, wages, and welfare overall might reduce. We do have a blueprint of consequences of severe trade conflicts and protectionist trade policies which is not even 90 years old.

To conclude, it is the author's belief that economists and liberals have the obligation to argue again in favour of the liberal world, and thus for free trade, to explain its mechanisms and to explicitly state policy recommendations to eventually convince their opponents and those that are currently indifferent. This doctoral thesis is meant to be a tiny contribution to this objective.

1 Introduction

Three chapters of this thesis deal with the effects of food standards on international trade. These build upon the gains of free trade as a fundamental conviction of economists. Thus, the overall question is to what extent and in which direction standards affect trade. If these cause protectionist effects, economists and policy makers need to take this trade-off between protectionism and ensuring food safety into account by designing standards as least trade-distorting as possible.

The fourth essay is unrelated to standards and trade. Instead, it deals with the most significant labour market reform – the Agenda 2010 – after World War II in Germany and how this affect unemployment and other outcome variables. It addresses a potential misperception that is often present in public debates. A causal inference is often mixed up with the simple correlation of law implementation during the years 2003 to 2005 and the continuous decline of the unemployment rate since 2005.

To relate the debate of standards and trade into the relevant context, the following Section 1.1 provides a short description of international trade and how it has evolved over time. Once the research context is sufficiently explained, the subsequent Section 1.2 introduces the reader to the debate of standards and trade which remains the core of the thesis. The final Section 1.3 summarises all four essays.

1.1 International trade flows are non-linear

Globalisation is often discussed as a natural consequence of technological progress following a one-way direction. However, historians (and economist who were fortunate enough to study economic history) tell us that globalisation, understood as increasing trade flows in this context, does not only intensify unboudedly but rather follows waves with many tiny peaks which are business-cycles. Many economist argue that the current integration of global markets is likely to slow down. For example, the *Economist* discusses the trend of shrinking multinationals which come back to their traditional home markets after intensive experiences of mergers and fragmentation of production processes abroad (Economist, 2017). This phenomenon has started earlier than the raising nationalism (briefly sketched in the preface). However, both factors might amplify each other since raising nationalism makes fragmentation more difficult and trade overall more costly. Thus, discussing the effects of standards on trade remains incomplete if this context of a possible slow-down of global trade is ignored because the raising relevance of standards is not the only new determinant of global trade flows.

Globalisation in terms of increasing international trade flows is not a new phenomenon. As a key element of the "long 19^{th} century" as well as of the "Belle Époque", the first wave of globalisation took place from mid 1800s until the begin of World War I in 1914. Falling transportation costs and technological progress as a result of the industrialisation facilitated global exchange of goods and factors of production. Nationalism, protectionist trade policies as a consequence of global recession in the early 1930s, and ultimately two worldwars impeded international trade and thus, stopped the first wave of globalisation (Helpman, 2011).

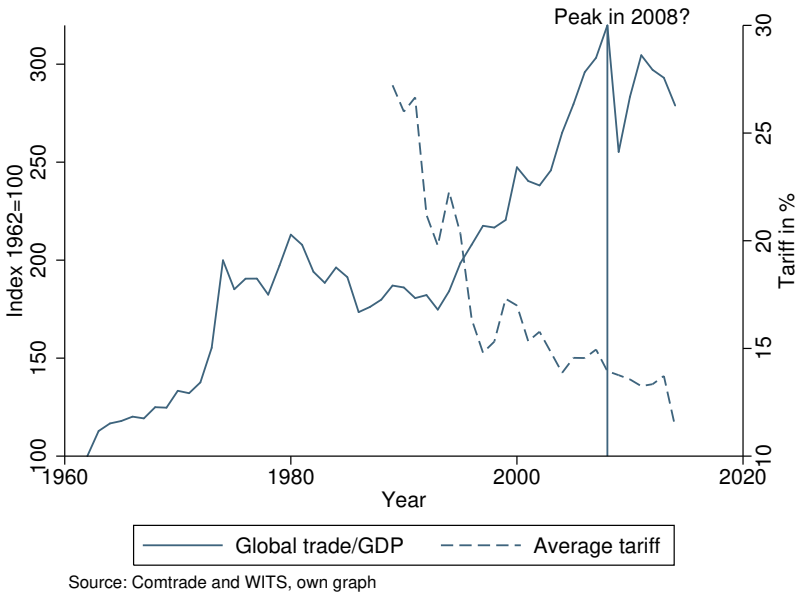
With the disasters of the 20^{th} century in mind, the world has been re-structured starting in 1944 and proceeding years. Global institutions like the United Nations, the World Bank, the International Monetary Fund, and the General Agreement on Tariffs and Trade (GATT) have been established to provide a regulatory framework for global exchange, financial stability, poverty reduction, and peace. The second wave of globalisation has been initiated.

Since then, the integration of global markets has intensified, international trade prospered, production processes became heavily fragmented, and many countries have experienced significant increases in per capita income and welfare. In the beginning, mostly European countries such as Germany, France, and UK as well as the USA, and Japan experienced high growth rates. In recent decades, many South East Asian countries such as South Korea, Singapore, Indonesia, but also India and China caught up and have grown significantly. Other countries such as Congo Kinshasa, Nicaragua, or Ivory Coast, however, have failed catching-up. Reasons for economic divergence are complex and not limited to trade and the degree of openness. The poverty-reducing influence of trade remains broad consensus among economists nevertheless (Frankel *et al.*, 1999; Dollar *et al.*, 2004). Thus, it remains of high interest which factors drive and which inhibit trade.

The increase of international trade flows has been about three times larger than the increase of global income. Figure 1.1 depicts the ratio of global trade flows and the gross domestic product (GDP) of the entire world with an index which is equal to 100 in 1962. Trade increased in the late 1970s and the early 2000s in particular. At the same time, tariffs (dashed line) have declined as a result of multinational negotiation rounds within the GATT-framework which became the WTO in 1994. This proxy of globalisation might has peaked in 2008. Scholars like Hoekman (2015) already proclaimed the "global trade slow down'" implying that the trade-GDP ratio not only took a break but indeed peaked in 2008. He argues that globalisation has an inherent limit because production processes can not be fragmented unboundedly. Fragmentation is one key driver of increasing international trade flows. If his view is true, the second wave of globalisation will slow down further.

Various factors influence global trade flows. Tariffs, quotas, export subsidies, and other trade policy measures have been used in recent decades for multiple objectives such as export support or protection of domestic producers. However, non-tariff barriers (NTBs) like norms and standards have gained in importance. Producers, consumers, and governments are increasingly interested in traded products that meet specific requirements regarding quality and processing.

FIGURE 1.1: The global trade slow down? (Hoekman, 2015)



Source: Comtrade and WITS, own graph

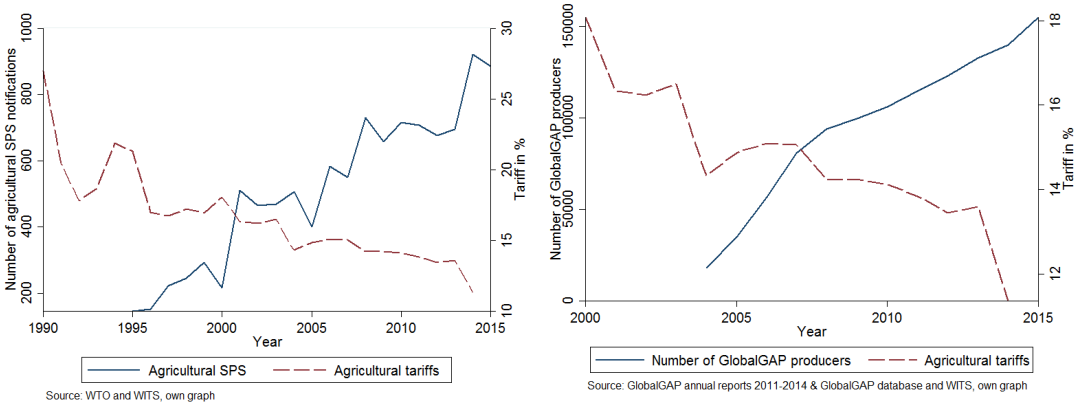
This thesis addresses the increasing relevance of standards as an additional determinant of trade. Standards constitute an important subset of NTBs. The effect of standards on trade needs to be investigated because standards could constitute a fundamental threat for international trade. If this is found to be true the raising relevance of standards would contribute to a declining trade-GDP ratio as depicted in Figure 1.1. The objective of the thesis is exactly the question whether standards constitute a barrier to international trade or whether standards could actually enhance trade flows. The following section provides the background and the current status of this debate.

1.2 Food standards as new determinants of global trade

The relevance of tariffs as trade barriers has declined during the recent decades while NTBs to trade have gained in quantitative as well as in qualitative importance (Hoekman *et al.*, 2011). For example, the total amount of sanitary and phytosanitary (SPS) notifications to the WTO as a proxy for public food safety standards increased from less than 200 in 1995 to almost 1,000 in 2015, see Figure 1.2. Moreover, the number of GlobalGAP producers as an important private standard increased from below 20,000 in 2004 to more than 150,000 in 2015 (*GlobalGAP Annual Reports*; Swinnen *et al.*, 2015). At the same time, agricultural tariffs have declined on average from about 27% in 1990 to 11% in 2015 (simple mean, all agricultural HS classifications, all countries). Therefore, the effect of standards on trade is of deep interest for economists and policy makers that are concerned about global trade and the integration of developing countries into the world trade system (Otsuki *et al.*, 1999).

Reasons for the increasing relevance of food standards are manifold. Traditionally, standards were set by governments to ensure food safety and to "prevent food adulteration and misbranding" (McClusky, 2007). However, consumers' preferences have also sharpened towards higher food quality in various dimensions (Batte *et al.*, 2007;

FIGURE 1.2: The raising relevance of standards



(a) Public: SPS notifications

(b) Private: Number of GlobalGAP producers

Swinnen, 2007). These are among others food safety, organic production, environmental concerns, sustainability as well as other objectives such as labour standards. Hence, standards address multiple objectives. Not only demand-side but also supply-side factors facilitate the increasing relevance of standards. Improved signaling either via certification or meeting of public standards first, reduces transaction costs within value chains and second, affects purchasing decision of the final consumer (Swinnen *et al.*, 2015). For example, lowering transaction costs within value chains belongs to the primary objective of the International Featured Standards (IFS) which is discussed in detail in Chapter 5. Among other goals, IFS aims to lower transaction costs by harmonisation of process standards. Furthermore, certified products are associated to be of higher value such that consumers have a higher willingness to pay (WTP). This allows firms to increase prices and hence, supplements incentives to comply with standards.

The landscape of existing food standards is heterogeneous such that food standards differ in various dimensions. First, the issuing organisation could be public governments or private organisations. For example, maximum residue limits (MRLs) are mostly set by governments to ensure that specific pesticide thresholds are not violated. If imported food products are found to exceed one particular threshold, the country can reject to import this product. In contrast, other standards are set privately. For example, the Fairtrade certificate is issued by Fairtrade International which is a private organisation located in Bonn and Cologne. Second, standards are either mandatory or voluntary. MRLs are mandatory because their violation allows countries to reject food products from being imported. In contrast, Fairtrade is a voluntary standard since no country requires their food imports to be Fairtrade-certified. Third, standards differ in terms of their position within value chains. Whereas Fairtrade traditionally addresses the final consumer, other standards like GlobalGAP are business-to-business (B2B) oriented. Finally, standards differ regarding their coverage. Whereas food safety was initially the primary objective of food standards, other objectives like environmental effects, nutrition requirements, and social concerns have gained in relevance as well (Swinnen *et al.*, 2015).

The implementation of food standards is legally covered by article 20 of the GATT

which allows countries to adjust their trade policies "to protect human, animal or plant life or health" as long as these are non-discriminatory. Based on this article, the WTO introduced the SPS measures which allow countries to set individual standards for this purpose as long as these are scientifically justified. The WTO also asks countries to make use of international standards to reduce potential trade barriers effects of standards. The Codex Alimentarius was initiated by the World Health Organisation (WHO) and the United Nations in 1963 to enable countries to harmonise their food standards. Nevertheless, many countries set standards individually to protect domestic consumers. However, the distinction between standards that are scientifically justified and standards that are primarily implemented due to protectionist motives often remains difficult to determine in practice. Therefore, it is of high interest to analyse to what extent standards enhance or reduce trade and under which conditions which effect predominates.

The effects of standards on trade remain ambiguous from a theoretical perspective. Lower information asymmetries, higher quality products, and lower transaction costs are expected to enhance trade (Jaffee *et al.*, 2005). In contrast, compliance with standards implies significant investment costs which are mostly of fixed costs but sometimes also of variable costs character. Compliance costs might exclude small-scale farmers in particular which would counteract development policies in developing countries. As a result, compliance costs could reduce trade. As we will argue in Chapter 4, compliance costs are not purely of monetary nature but also have a significant non-monetary part.

Empirical analyses on the effects of food standards on trade flows remain unclear as well. Previous studies found empirical support for trade-enhancing as well as for trade-reducing effects. Jaffee *et al.* (2004) entitled this debate as "standards-as-barriers to trade vs. standards-as-catalysts to trade". As one of the main references, Otsuki *et al.* (2001) analyse the effect of the EU's new harmonised aflatoxin standard on food products imported from a set of African countries by estimating a gravity equation. They find that African exports of edible groundnuts and groundnut oil are constrained by MRLs of aflatoxin set by EU member countries during 1989 and 1998. Xiong *et al.* (2012) complements the analysis of Otsuki *et al.* (2001) by using a panel dataset instead of cross-section and a different method which accounts for sample selection bias. The availability of panel data allows to account for time-invariant heterogeneity and multilateral resistance. Sample-selection might cause biased estimates because taking the logarithm of the depended variable eliminates all zero trade flows. Hence, Xiong *et al.* (2012) follow a two-step procedure based on Heckman (1979) and Helpman *et al.* (2008). In addition, ignoring zeros does not allow the study to estimate the extensive margin of trade, that is the creation of new bilateral trade partnerships. In other words, all estimates are based on the assumption that trade already takes place; no conclusions can be made on any implications for new trade flows. The objective of Xiong *et al.* (2012) is to examine the effect of the harmonisation and tightening of EU MRLs on aflatoxins in 2002 on African exports of three groundnut products by accounting for sample selection bias, multilateral resistance, and the heterogeneity across firms. They find that EU MRLs have no significant influence on groundnut exports from Africa across all preferred methods of estimation. Domestic supply conditions in Africa are crucial in order to determine trade volumes and the propensity to trade groundnut products. Roy (2013) adds an additional aspect by arguing that African exports of groundnut had already declined before the

harmonisation by the EU took place.

The relevance of the chosen method is also highlighted by Ferro *et al.* (2015) who create a restrictiveness index of MRL pesticide for 61 importing countries. By applying the two-step Heckman procedure as illustrated by Helpman *et al.* (2008), the authors find evidence in the first stage that more stringent MRLs reduce the probability to export due to higher fixed costs. However, once the sample selection bias and the share of exporting firms are controlled for, standards have no effect on trade flows. In addition, the first-stage effect is stronger for BRIC-countries than for non-BRIC countries. Exports from low income countries are more negatively affected by product standards than those from higher income countries. Ceteris paribus, countries export to destination markets which have the lowest fixed costs, i.e. less restrictive MRL standards. The effect of food safety standards on China's exports is also analysed by Chen (2008) who find a statistically significant negative effect. According to their estimates, the effect is even stronger than imposing tariffs. Further evidence for trade-reducing effects due to more restrictive standards is – among others – also provided by Wilson *et al.* (2003), Chen *et al.* (2006), Yue *et al.* (2010), Drogué *et al.* (2012), and Melo *et al.* (2014) who all focus on the effects of MRL on exports.

A meta-analysis of Li *et al.* (2012) aims to draw some general conclusions out of the broad empirical literature. By using the available estimates, the authors find that the agricultural sector compared to any other sector is more characterised by the barriers-to-trade view. Moreover, not addressing empirical challenges of the gravity framework such as multilateral resistance increases the likelihood of finding negative effects of standards on trade flows. Similarly, using analyses on MRLs are more likely to find trade impeding effects.

In a novel approach Xiong *et al.* (2014) address both theoretical arguments of trade impeding and trade enhancing effects of trade flows. They disentangle the effect and find MRLs to increase import demand and reduce supply of exporting countries. Again, developing countries are negatively affected.

Cost of compliance with food standards remain the major reason causing trade-reducing effects due to exclusion of farmers from global value chains. (Kleinwechter *et al.*, 2006; Schuster *et al.*, 2015; Kiss *et al.*, 2003; Colen *et al.*, 2012). Compliance costs occur – for example – because of investments in specific production facilities or new requirements regarding production processes. However, standards could also act as non-financial barriers in the sense that non-financial resources are required to comply (see Chapter 4). Compliance costs amplify once each country sets its individual standard. Maskus *et al.* (2005) use firm-level data of 159 firms located in 16 developing countries of the World Bank Technical Barriers to Trade Survey Database to estimate fixed costs of compliance. They find those to equal $425,000 per firm on average which is about 4.7% of value added on average. Moreover, a one-percent increase of investment costs of standards in importing countries increases variable production costs *c.p.* on average between 0.06% and 0.13%. Based on the same analysis but emphasised by Czubala *et al.* (2009), these fixed costs vary by region. Firms in Sub-Saharan Africa face fixed costs about 7.65% of firm sales, firms located in the Middle East about 6.67%, and firms in South Asia only 1.79%. Henson *et al.* (2004) estimate an increase of production costs for Nile perch exporters in Kenya as a direct

result of stricter food safety standards of about 25%. As another example, investment costs due to implementing Hazard Analysis Critical Control Points (HACCP) and other procedures sum up to about $40,000 per plant. All in all, there is general consensus about the economic significance of fixed costs of compliance as relevant NTB.

Whereas most empirical studies emphasise the potential trade impeding effect of more stringent or heterogeneous food standards, Henson *et al.* (2008) highlight the opportunities associated with standards for developing countries. Without denying possible trade reducing effects, they argue that developing countries can use standards of developed countries to re-locate themselves on global export markets by particularly addressing preferences for these standards. However, the authors admit that firms' abilities to re-locate might be concentrated among the leading exporters whereas other – i.e. smaller firms – exporters are likely to be excluded. Research questions of Chapter 3 address the heterogeneous effects between leading exporters and the rest explicitly.

Empirical studies on private standards are less frequent probably because relevant data are more difficult to access. However, these are at least equally important because private standards gained in relevance in recent years in particular. Furthermore, Fulponi (2006) shows that private standards are on average stricter than public standards. Most private standards are considered as *de facto* mandatory although they remain legally voluntary (Henson *et al.*, 2010). Schuster *et al.* (2015) focus on the effect of a broad set of private standards on exports of Peruvian asparagus firms. Based on a rich dataset including 18 years and 87 firms, they do not find an effect. In contrast, Masood *et al.* (2014) find that GlobalGAP certification increases banana exports to the EU. Again, empirical results on trade are mixed for private standards as well as for public standards.

Standards affect economic development of countries not only via changes in trade flows. For example, Colen *et al.* (2012) find empirical evidence that GlobalGAP certification in Senegal improves employment conditions. Chiputwa *et al.* (2015) show that Fairtrade certification increases income of coffee farmers in Uganda. Moreover, standards could help developing countries to meet quality requirements of importing countries. Low productivity and low quality are the most relevant barriers for developing countries to enter global value chains. While a lack of productivity can be offset by low wage levels, improving quality is more challenging (Swinnen *et al.*, 2015). Because compliance with standards often signals a specific quality of particular products, the proliferation of standards might be a suitable development strategy for developing countries to integrate into global value chains. For example, several private standards like GlobalGAP, TESCO, IFS, and British Retail Consortium (BRC) facilitate vertical integration of Peruvian asparagus farmers (Schuster *et al.*, 2013). Table 1.1 provides an incomplete overview by listing a set of related studies which the author considers to be relevant contributions in the field.

To conclude, the literature review reveals a large set of studies with diverse outcomes. Thus, no direct one-suits-all answer exist such that deeper analyses are required which detect under which conditions the catalyst-view or the barriers-view predominates.

This thesis contributes to the ongoing debate by analysing two different food standards in econometric settings reflecting the current research frontier. These standards are first, MRLs as an important public standard (Chapter 3) and second, IFS which is an important private standard (Chapter 5). The following section summarises all four essays briefly. All essays build upon a flexible conceptual framework based on Melitz (2003) which is explained in Chapter 2.

1.3 Summaries of essays

The first three essays (Chapters 3, 4, and 5) analyse under which conditions positive or negative effects of standards on trade overweight. We look at the following different determinants: market structure, product complexity, quality of public institutions, and type of standards. The first essay addresses the role of market structures by arguing that country-pairs with high trade volumes will trade even more once standards become stricter whereas country-pairs with low trade volumes trade even less. The second essay uses fixed costs of compliance as well as point of departure. It focuses on non-monetary aspects of fixed costs by analysing the role of institutional quality and its relevance for exports of various food products which differ in terms of complexity and degree of processing. Finally, the third essay is complementary to the second essay since it argues that product complexity matters for the direction of the effect of standards on trade. It considers the IFS as an important private standard for processed food products.

The fourth essay in Chapter 6 is unrelated to the overall topic of food standards and trade. It focuses on the effects of the "Agenda 2010" as the most comprehensive labour market reform in Germany after World War II on labour market outcomes. This chapter is a novel contribution for this field of literature because it applies synthetic control methods (SCM) as unique causal identification strategy. Moreover, these reforms were meant to make the rigid German labour market more flexible and were accompanied by intensive debates in public. Thus, this chapter is of high policy relevance.

Essay one: Exporter size matters – heterogeneous effects of food standards on trade flows

With increasing restrictiveness of standards, leading exporting firms are found to benefit in terms of trade volume while other firms are disadvantaged (Anders *et al.*, 2009). Thus, the *ex-ante* export volume seems to be relevant for the effect of increasing standards on producers. Leading exporters might be more capable to cover the fixed costs of compliance while such an investment will be less profitable if expected exports are relatively low. Consequently, we expect the effect of food standards on trade to depend on the *ex-ante* export volume.

We argue that this phenomenon is not only present at the firm-level but also at the country-level because agricultural export markets are highly concentrated. Few exporting countries often account for a large share of world exports of a particular

TABLE 1.1: An overview of related studies – additional information indicated by the letters in superscript is provided in Table A.1.

Authors	**Exporter**	**Importer**	**Products**	**Years**	**Type of Standard**	**Method**	**Effect**
Otsuki *et al.* (2001)	Africa[a]	EU-15[b]	Cereals, dried fruits, nuts[c]	1989-1998	EU aflatoxin	FE	negative
Xiong *et al.* (2012)	Africa	EU-15	Groundnuts[d]	1989-2006	EU aflatoxin	Trun-OLS, PPML, NBPML ZIPPML, ZINBPML, HMR	no
Ferro *et al.* (2013)	World?	61	66	2006-2011	MRL	OLS HMR	negative no
Anders *et al.* (2009)	33	US	Seafood	1990-2004	HACCP	RE	negative/ positive[e]
Chen (2008)	China	9-10[f]	Vegetables Aquatic products[g]	1992-2004	Chlorpyrifos MRL Oxytetracycline MRL	OLS & Imp-FE	negative
Chen *et al.* (2006)	17[h]	5[i]	?	?	WB TBT	GLM & Exp./Ind.-FE	negative
Disdier *et al.* (2010)	?	EU, Canada US, Japan	Crustacean	2001-2006	Chloram- phenicol MRL	OLS HMR	positive no
Drogué *et al.* (2012)	7[j]	7[k]	Apples Pears[m]	2000-2008	MRL	NBR, ZINB	mixed[l]
Melo *et al.* (2014)	Chile	15[n]	Grapes, apples kiwi, cherries	2005-2009	Aggregate Disagg.[o]	OLS, PPML	negative no[p]
Scheepers *et al.* (2007)	South Africa	EU	Avocado	?	Prochloraz MRL	POLS	negative
Schuster.2013a	Peru	EU	Asparagus	1993-2011	Private standard	OLS FE sys. GMM	no no no
Wilson *et al.* (2004)	21[q]	11[r]	Banana[s]	1997-1999	Chlorpyrifos MRL	FE	negative
Winchester *et al.* (2012)	World?	10[t]	11[u]	2008-2009	MRL other[v]	PPML, GNegB, ZIP	negative no
Xiong *et al.* (2013)	61	20[w]	109[x]	2008, 2012	MRL	HMR	no[y]
Yue *et al.* (2010)	8[z]	EU-14[a]*a*	Tea	1997-2006	MRL	OLS & Imp.-FE	negative

product. The resulting hypothesis is that the catalyst-view predominates for the leading exporting countries while the barriers-view applies for country-pairs which trade relatively little. Hence, the novelty of this chapter is to contribute to a solution of the debate of barriers- vs. catalyst views by providing insights when which view is applicable.

Because the gravity model of trade uses bilateral trade flows as dependent variable, quantile regressions are our preferred method to analyse the hypothesis of heterogeneous effects of standards on trade. We employ a rich data set including MRLs as our target variable covering ten years from 2005 until 2014. MRLs are an important public food safety standards which have been frequently studied in previous works. This makes our results comparable to earlier findings.

We apply quantile regressions within the gravity framework by following a three-step procedure which addresses empirical challenges that became a matter of concern in recent years. Our empirical strategy builds upon Figueiredo *et al.* (2014) who successfully control for multilateral resistance, selection-bias, and the high share of zeros. Moreover, we conduct robustness checks regarding construction of the dataset.

As a result, we indeed find empirical support for the hypothesis. To be more precise, standards impede bilateral trade between country-pairs from the first to the eighth decile while country-pairs at the highest decile trade even more as standards become stricter. We also find an heterogeneous effect of tariffs on trade such that we devote analyses of heterogeneous effects of trade costs overall on trade to future research. Thus, the contribution of the chapter is that the effect of standards on trade is not only determined by income but also by the relative export volume. This is a relevant finding because previous research mostly focused on income as a reason for non-compliance. In the long run, these results imply that stricter standards could increase the degree of concentration of agricultural export markets.

Essay two: Institutional quality and the ability of organisations to adopt foreign agri-food trade standards

Surveys conducted with farmers reveal that non only income-constraints inhibits compliance with standards but also non-monetary factors. We construct a conceptual framework in which the standard-adoption-ability (SAA) at least partly determines whether an organisation complies or not. We argue that non-monetary aspects like the institutional setting affects SAA. Moreover, we claim that SAA is in particular important for compliance with food safety standards of rather complex products.

To test the hypotheses, we use the gravity model in which production dummies interacted with the quality of public institutions and exports to the EU constitute the main variable of interest. This empirical strategy relies on the fact that the EU belongs to the strictest importers of agricultural products. Furthermore, we use five different agricultural product categories which are dairy, meat, fruits, vegetables, as well as fish products. Due to their animal origin, we consider dairy, meat, and fish products to be more complex to produce. Public institutions of high quality are likely

to support farmers to produce relatively complex products. Thus, we expect the effect of institutions of the exporting country on trade to be more pronounced for these products and less important for exports of fruits and vegetables.

The Worldwide Governance Indicators (WGI) are used as proxies for the quality of public institutions (Kaufmann *et al.*, 2010). These are six indicators which have been calculated based on firm level surveys and other indices. WGI aim to reflect the functionality of a state such as rule of law or effective governance. Principal component analysis is applied to comprise all six indicators to one single variable.

We find empirical support for the hypothesis for all products except fish. The quality of institutions is particularly important for exports of dairy and meat products but not relevant for the other three categories. We substantiate the results by providing descriptive data on the number of exporting countries for each product category. The number of countries which export meat and dairy products is much smaller than for the other products in absolute in relative terms. We interpret this as additional support for our hypothesis since it indicates – but not proves because production decisions of agricultural products obviously depend on other conditions like climate and soil characteristics as well – that relatively few countries are capable to export dairy- and meat products to the EU.

Essay three: the role of private standards for manufactured food exports from developing countries

The third essay focuses on exports of processed food products instead of agricultural raw products. We estimate the effect of IFS certification as an important private standard on agricultural exports. IFS certifies processed food products at upper parts of value chains. This feature makes IFS distinct from other private standards like GlobalGAP which is more frequently analysed. IFS is a post-farm gate standard whereas GlobalGAP is a pre-farm gate standard mainly for agricultural raw products that originate directly from farms. The focus on processed food items in contrast to raw products is a particular relevant subject to analyse because the value added of production is higher for processed food products than for raw products. In addition, IFS explicitly aims to reduce transaction costs within the value chain which might support the trade-catalyst perspective.

As key novelties of this chapter, we have data of IFS certification for six years of 12,000 producers located in 88 different countries. Furthermore, we use certification in neighbouring countries of an exporting country as an instrument to address endogeneity due to reverse causality. Finally, we conduct the analysis by income of the exporting country and at the sectoral level.

As a result, we find that IFS certification *c.p.* increases exports on average. However, once we separate exporting countries by income, the positive effect only remains for high-income countries while it turns negative or statistically insignificant for upper-middle and low-income countries respectively. This finding reveals important policy implications because the discussed raising relevance of standards then impose an additional burden on low-income countries. Private standards are not a one-suits-all

development tool to increase agricultural exports of developing countries. Instead, our results stay in contrast with other studies which emphasise the potential of GlobalGAP certification to integrate Senegalese farmers into global value chains (Colen *et al.*, 2012).

Essay four: the Hartz reforms and the German labour Force

The fourth essay in Chapter 6 of this thesis investigates whether the most comprehensive labour market reform "Agenda 2010" in Germany after World War II caused the decline of the unemployment rate in 2005 and subsequent years. There is little empirical evidence based on causal identification strategies that support this conclusion. This is surprising because there seems to be broad consensus in public debates in Germany that these reforms were responsible for the decline of the unemployment rate and for the current economic strength of Germany despite the crises in Europe. Therefore, we consider this research question to be highly relevant for policy makers and society overall.

The Agenda 2010 and the Hartz-reforms I to IV were subsequently implemented during the years 2003 to 2005. These reduce the amount and duration of unemployment benefits. Moreover, they facilitate temporary employment and tighten conditions when job-seekers need to accept a job offer. These reforms were debated intensively in Germany and ultimately forced chancellor Schröder (social-democratic party) to resign. The long-run effects for the political landscape are enormous. A new left-wing party was founded and is well established in German parliaments. The responsible social-democratic party is still divided between opponents and advocates of these reforms. Moreover, research on the effect of these reforms affects policy implications for economic policies of European countries whose economic conditions are more worrisome than Germany's.

We apply a novel causal identification strategy which relies on SCMs. The underlying idea is to construct a counterfactual which becomes the control group. This synthetic Germany is built by using a donour pool of OECD countries with a broad set of predictor variables like schooling, investments, demography, and many other explanatory variables. We do not only use unemployment as outcome variable but also labour force participation (LFP) and employment of working age population. Moreover, we distinguish by gender and age groups. As a result, the effect of the Hartz reforms on unemployment remains weak. But the reforms have increased LFP, especially of women and elderly people.

2 Conceptual framework

This chapter introduces the conceptual framework of the first three essays and of Chapters 3 and 5 in particular. It presents aspects of a model based on firm heterogeneity following Melitz (2003) and quality upgrading. However, it is not meant as a complete and full theoretical model which is eventually brought to data. Instead, it should sketch the lines of thought of the empirical strategies that follows in Chapters 3, 4, and 5.

2.1 Aspects of the Melitz-model

The essential role of fixed costs for production and exports has been emphasised in the "New-trade-theory" as well as in the "New-new-trade theory". Whereas the former is mostly motivated to explain intra-industry trade by implementing product differentiation in a monopolistic competition framework, the latter relaxes the assumption of firm homogeneity by arguing that exporting firms have fundamentally different characteristics than non-exporting firms in terms of productivity, wages, production volumes, and profits. We use elements of the "New-new-trade theory" of the Melitz model to demonstrate the effect of stricter food standards on bilateral exports (Melitz, 2003).

Melitz introduces firm heterogeneity via the productivity parameter φ. Firms need to pay sunk entry costs f_E to draw their productivity level from a cumulative Pareto distribution $G(\varphi)$. This productivity level determines whether the firm exits the market, serves the domestic market only, or even exports to foreign markets. Export costs from country i to country j are denoted by f_{ij}. These include costs for production, exports, and market access. Thus, if a firm sells to the domestic market only, f_{ij} is denoted as f_{ii}. Hence, profits π_{ij} of all producing firms are given by Equation 2.1:

$$\pi_{ij}(\varphi) = B_j \tau^{1-\sigma} \varphi^{\sigma-1} - f_{ij} \tag{2.1}$$

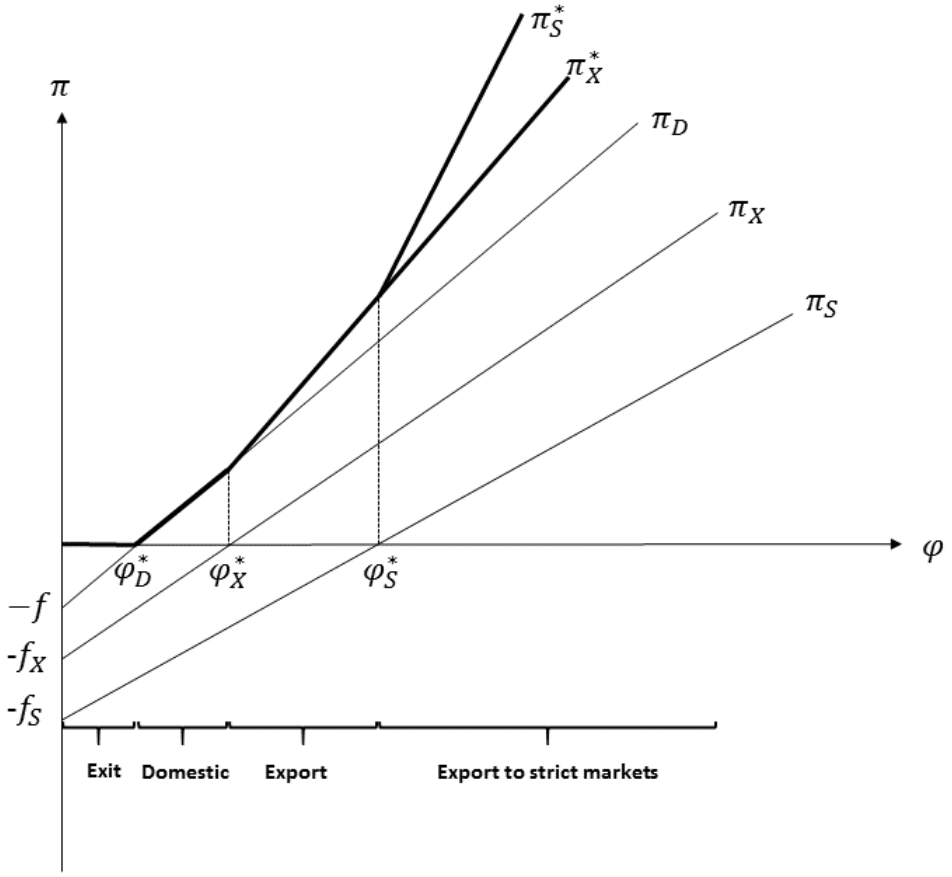
B_j is a demand parameter of the destination market j, τ represents iceberg-type trade costs, and φ is the drawn productivity parameter. σ denotes the constant elasticity of substitution which is assumed to be greater than one. Thus, the zero-profit cutoff condition for exporting from market i to market j yields the cutoff productivity level φ^*_{ij} at which profits are zero, see Equation 2.2.

$$\pi_{ij}(\varphi^*_{ij}) = 0$$

$$\iff \quad B_j(\tau_{ij})^{1-\sigma}(\varphi_{ij}^*)^{\sigma-1} = f_{ij} \tag{2.2}$$

In equilibrium, higher fixed costs f_{ij} are associated with higher demand, lower trade costs, or higher productivity for $\sigma > 1$. Figure 2.1 depicts the relation between fixed costs and various cutoff productivity levels. If the drawn productivity level is below φ_D^*, the firm exits the market; if $\varphi_D^* < \varphi < \varphi_X^*$, the firm produces for the domestic market only but does not export. Once the productivity level exceeds φ_X^*, the firm exports. Note that the slope of the corresponding profit curve π_X is smaller than for π_D due to variable trade costs. Profits for exporting firms are jointly determined by π_D and π_X and given by the bold curve π_X^*.

FIGURE 2.1: Profits, productivity, and standards in the Melitz framework.



2.2 New-new-trade theory and food standards

As argued in the introduction, compliance with food standards requires additional fixed costs. Melitz already defines fixed costs broadly as "market access" costs. Therefore, the stricter food standards in the destination market become, the larger are market access costs. Although compliance with food standards might also increase variable costs, e.g. due to more intensive auditing, the implied fixed costs because of investment in modern production technologies are considered as potential barriers

to trade. Therefore, we add a firm-specific quality upgrading fixed cost term $f(q_i)$ where q_i is a firm-specific quality parameter for differentiated goods.

However, we implement food standards not only via increased fixed costs of exporting at the supply side but also as a strategy to address preferences of modern consumers at the demand side and hence, as a form of quality upgrading (Ferguson, 2009). Thus, profits also increase in q_i. The zero-profit cutoff condition 2.2 then changes to:

$$\max_{q_i} \quad \left[q_i B_j \left(\tau_{ij}\right)^{1-\sigma} \varphi^{\sigma-1} - f(q_i) - f_{ij}\right] \tag{2.3}$$

Hence, the optimal upgrading choice is given by:

$$B_j \left(\tau_{ij}\right)^{1-\sigma} \varphi^{\sigma-1} = \frac{\partial f(q_i)}{\partial q_i} \tag{2.4}$$

Thereby, we assume that the firm-specific quality upgrading fixed costs $f(q_i)$ are continuously differentiable which implies that conformance with a specific quality-level, and hence standards-requirements, is not a binary decision. Instead, an optimal standard can be chosen from a broad continuum of standards. Following Ferguson (2009), we need to specify the functional form of $f(q_i)$. We assume that meeting relatively low levels of standards is a low hanging fruit. However, costs are expected to increase exponentially since it becomes increasingly difficult to meet high levels of standards. Thus, we assume that quality upgrading fixed costs are convex and increase in q_i. Hence, the partial derivative of $f(q_i)$ with respect to q_i increases in q_i. Higher levels of standards imply an increase of the right-hand side of Equation 2.4 which in equilibrium requires *c.p.* higher levels of demand B_j.

Krugman's "new-trade theory" introduces production technologies which are often characterised by increasing returns to scale due to the fixed costs of production. Hence, larger production volumes result in lower average costs. In other words, if the demand B_j of a firm is *ex-ante* relatively small, fixed production costs in the zero-profit cutoff conditions 2.2 and 2.4 become binding at relatively low levels. Consequently, firms that face lower demand are more likely to exit the market – because they need to draw *c.p.* larger productivity parameters – if fixed costs of production and market access costs increase due to stricter standards. Given the same drawn productivity level, market demand B_j determines whether firms exit the market as f_X increases to f_S. Hence, the level of demand could determine compliance with standards or not. This completes our conceptual framework for Chapter 3.

However, the model also allows to address the trade-off between quality-upgrading (compliance with standards) to earn higher profits and non-compliance. It relies on the idea that firms can either improve quality of their products to earn higher profits or to export to other markets which do not require products of such quality. A case-study of Kenyan exports of Nile perch illustrates the trade-off (Henson *et al.*, 2004). The authors show that only a subset of all Kenyan exporters upgraded production

capacities to meet for example stricter hygiene standards while others started to export to countries with lower standards.

To be more explicit, we assume the following functional form of $f(q_i)$:

$$f(q) = q^{1/\theta} \quad \text{with} \quad \theta \, \varepsilon \, [0, 1] \tag{2.5}$$

The shape-parameter θ indicates the "ease" of quality upgrading (Ferguson, 2009, p.10). The larger θ is, the easier a firm can address preferences of consumers that demand high-quality products; i.e. products that meet relatively strict standards. Hence, if firms are able to implement standards easily, they will benefit from a lower increase in associated costs which – eventually – require lower levels of demand.

Using this specific functional form of the costs of quality-upgrading, the optimal quality-level is then given by:

$$q^* = \left(\theta B_j \tau_{ij}^{1-\sigma} \varphi^{\sigma-1}\right)^{\frac{\theta}{1-\theta}} \tag{2.6}$$

Keeping B_j, τ_{ij}, and φ constant, higher values of θ increase the optimal level of quality. Thus, if a standard is particularly capable to address consumers' preferences – i.e. high values of θ – producers can earn higher profits by investing in stricter – i.e. high-quality – standards. In this scenario, we would expect standards to increase profits and therefore, increase trade flows at the aggregate level. Contrary, if a particular standard is less capable to address consumers' preferences – θ is close to zero – the quality level remains low and producers are less likely to invest in the standard.

Therefore, it is the nature of θ that determines whether firms comply with standards of importing countries or not. Thus, θ captures the characteristics of the specific standard, which in-turn determine whether trade flows increase or decline in the corresponding context.

As we will argue, MRLs as standards in Chapter 3 are assumed to have lower values of θ because their primary objective is not to address consumers' preferences to generate higher profits. Instead, MRLs are set by governments or the WTO to ensure food safety. Therefore, we expect an overall trade-reducing effect of MRLs. In contrast, IFS as an important private food standards, are introduced by major retailers to reduce transaction costs. Their objective is to address consumers' prefereces well such that we assume higher values of θ and as a result, increasing trade flows due to higher profits.

3 Exporter size matters – heterogeneous effects of food standards on agricultural trade

Abstract:

The effects of food standards on bilateral trade flows remain unclear. Trade theory provides arguments for trade enhancing as well as for trade impeding effects. According to previous studies, standards could exclude developing countries in particular from high value trade chains due to the high fixed costs of compliance. However, standards also reduce information asymmetries and address changing preferences of modern consumers. Using descriptive evidence for highly concentrated agricultural export markets as our point of departure, we expect the effect of food standards to differ for leading exporting countries compared to those which export relatively low volumes. Based on high fixed costs of compliance in combination with increasing economies of scale, we expect leading exporting countries to be more likely to invest in food safety standards whereas countries with lower export volumes are not capable investing. Hence, we claim that it is not only the income level of the exporting country that determines whether compliance is feasible but also the size of exports. The application of a three-step quantile regression procedure within the gravity framework builds the empirical core as this method accounts for heterogeneous covariate effects along the trade flow distribution. Using MRLs of 54 countries and 50 products from 2005 to 2014, we find the expected heterogeneous effects of MRLs on agricultural trade at the aggregate level. We conclude tat the previously ignored *ex-ante* export volume determines how food standards affect trade.

This chapter is joint work with Bernhard Brümmer and Inmaculada Martínez-Zarzoso. Malte Ehrich's contributions are: Development of research question and research design, empirical and theoretical framework, data management, empirical estimation, and writing of the manuscript. All authors reviewed and edited the manuscript.

Acknowledgements: We received valuable help for the Stata code from Prof. Erik Figueiredo. Furthermore, we are grateful to our research student assistants Sascha Berndt and Dela-Dem Doe Fiankor for excellent support in data cleaning and the descriptive analysis.

3.1 Introduction

Agricultural export markets are characterised by a high degree of market concentration. For example, the five leading exporting countries accounted for more than 90% of world exports of almonds in 2014; the share for apples was about 63%, for bananas around 70%, and for garlic about 92%. Table 3.1 shows these export shares for 48 different agricultural products. With the exception of tobacco, the export shares of the five leading exporting countries of all listed products exceed 50%. We expect this pattern to be an important determinant of trade flows. As we will argue, the degree of market concentration is in particular relevant – and currently missing – in the discussion of "standards-as-barriers vs. standards-as-catalysts" to trade, see Section 1.2.

Market concentration becomes relevant in this context because of compliance costs of food standards. Their fixed costs character – either monetary or non-monetary, see Chapter 4 – might prevent countries with low export volumes of particular products to enter global agricultural export markets. Thus, we expect countries with low export volumes to be less likely to invest in food standards. These countries have lower economies of scale to cover fixed costs. As a result, we expect the effect of standards on exports to differ along the trade flow distribution. Country-pairs with relatively low trade volumes might trade even less once standards become stricter whereas country-pairs at the upper part of the trade flow distribution might trade even more. If we find such heterogeneous effects of standards on trade, this might reconcile competing results within the literature. Thus, neither the catalyst-view nor the barriers-view always predominates. Instead, both are correct depending on the *ex ante* level of trade.

TABLE 3.1: Concentration-ratio of various agricultural products defined as: sum of export shares $\sum_{i=1}^{5} X_p^i / X_p^{world}$ of five leading exporting countries on world exports in 2014. *Source: UN Comtrade, author's calculations*

Product	**Export share (%)**	**Product**	**Export share (%)**	**Product**	**Export share (%)**
Almonds	97.57	Grapes	56.79	Pineapple	72.97
Apples	62.20	Hazelnuts	90.49	Plums	64.82
Apricots	74.86	Hop	89.79	Potatoes	57.32
Asparagus	84.07	Kiwi	87.76	Raspberries	81.51
Aubergine	75.64	Lettuce	77.67	Rice	80.18
Avocados	81.14	Maize	73.65	Rye	76.51
Bananas	66.02	Mandarins	75.25	Sorghum	95.16
Barley	64.87	Mangos	53.41	Soybeans	93.29
Broccoli	75.66	Millet	99.93	Spinach	85.85
Carrots	62.03	Oats	83.26	Strawberries	79.37
Celery	82.39	Olives	87.79	Tea	66.37
Cherries	100	Onions	62.61	Tobacco	37.35
Chestnuts	100	Oranges	69.77	Tomatoes	65.28
Cucumbers	77.99	Papayas	74.75	Walnuts	91.34
Garlic	92.30	Peaches	71.97	Watermelons	72.44
Grapefruit	71.27	Pears	99.95	Wheat	65.23

Krugman (1979) emphasises the role of market size and market concentration for international trade. He introduces the concept of fixed costs and increasing returns to scale within a monopolistic competition framework. According to his theory, the absolute number of firms worldwide that produce a specific product (or variety) reduces once countries switch from autarky to frictionless trade. We argue this mechanism is

more pronounced the higher fixed costs of trade are. Since food standards increase fixed costs of exporting, the increasing relevance of standards is likely to enhance this mechanism identified by Krugman (1979).

We base our hypothesis on the conceptual framework explained in Chapter 2. However, instead of focusing on productivity itself our framework emphasises on the role of demand as an additional determinant of profits and trade. If demand in the importing country is high, the exporting country is more likely to invest in standards because the investment is more likely to pay off. To put it differently, we claim that the existing trade volume determines whether standards act as barriers-to-trade (low *ex-ante* trade volume) or as catalyst-to-trade (high *ex-ante* trade volume).

Literature related to the role of exporter size

Anders *et al.* (2009) provide empirical evidence that the introduction of HACCP increases exports of leading exporters of seafood and reduces exports of countries with lower exports of seafood. Moreover, developing countries are more likely to experience lower exports as a response to stricter standards than developed countries. As a drawback, this article is limited to seafood products. It also does not include standards as a separate explanatory variable; instead standards are represented by a time-dummy which indicates the introduction of the HACCP. The authors apply random effects which is often found to be a non-valid method because of the likely correlation between omitted time-invariant heterogeneity and the error term. Thus, future research is required to further validate their findings.

The relevance of "historical" trade for current trade flows has been shown a while ago by Eichengreen *et al.* (1998). By including the one-year lag of the dependent variable, which is trade, into the gravity framework, the authors argue in favour of historical trade as an important determinant of current trade flows. This result is – together with the high concentration of export markets and compliance costs – one essential cornerstone for the argumentation of this chapter. If *ex ante* trade flows between two countries are large, they are not only more likely to trade more *ex post*, but also – as we show – more likely to comply with standards.

The study which bears the most similarity to our own analysis was conducted by the World Bank (Fernandes *et al.*, 2015). The authors apply a rich firm-level dataset which allows them to analyse the effect of relative restrictiveness based on MRLs on firm level exports. Among other results, the authors find that firms' past export volumes increase the probability of a firm to enter or to stay in the market. Smaller exporting firms are less likely to enter new markets as standards increase. Hence, this analysis addresses a similar research question at the firm level. We aim to provide further evidence for the relevance of export volume at the country-level with a different method. Herzfeld *et al.* (2011) analyse compliance of firms in developing countries with the BRC Food Technical Standards and GlobalGAP as two important B2B private standards. The authors show that both standards act as additional market entry barriers and that country pairs with well established trade relationships are more likely to comply with these standards. Graffham *et al.* (2007) demonstrate

that large scale producers in particular are more likely to comply with EuroGAP[1]. These findings already indicate the importance of production volume and trade volume for compliance with standards. We contribute to these findings by applying quantile regressions as a novel method in this field combined with a rich dataset at the country-level.

By using data of Vietnamese pangasius farmers Hansen *et al.* (2014) applied quantile regressions in order to analyse the influence of standards on income of relatively poor farmers compared to relatively rich farmers. As a result the authors find that only farmers between the 50% and 85% quantiles benefited from standards in terms of per capita consumption expenditures. Poor farmers were excluded and the income effect for the very rich farmers was negligible.

If trade volumes determine whether exporting countries comply with standards of the importing country or not, we expect a negative effect of standards on trade flows of country pairs with small trade flows and a positive effect (or less negative effect) for countries with relatively high trade flows. Quantile regressions allow to estimate coefficients for various quantiles and are therefore, a suitable method to test our hypothesis. As a result, quantile regressions allow for a richer interpretation than standard regression analysis which are narrowed to the conditional mean.

In addition to the first novelty of analysing heterogeneous effects via quantile regressions, we consider our dataset as a second novelty. We employ a rich dataset which contains trade flows between 54 importing and exporting countries of 49 agricultural raw products from 2005 to 2014 summing up to 1,402,380 observations. We use data on MRLs which define important public food safety standards. Countries define maximum levels of tolerated residues in agricultural products to minimise the risk of unhealthy food consumption.

The remainder of the chapter is organised as follows: Section 3.2 introduces quantile regressions in the context of gravity estimations. Because of a broad range of advantages the gravity model has been established as the workhorse model in empirical trade analysis. We explain how quantile regressions can be embedded into the gravity framework and how major empirical challenges of gravity models can be addressed, in particular heteroskedasticity, the large share of zeros in trade data, and the high number of fixed effects. Section 3.3 reports results, which are then discussed in the subsequent Section 3.4. A conclusion summarises the chapter.

3.2 Methods and data

We now link the conceptual framework of Chapter 2 with our empirical model. We are aware of the fact that the traditional Melitz-model is primarily designed at the firm level. However, in principle the underlying mechanisms and channels of the model are applicable to the country-level as well. This can be visualised by thinking of countries that draw productivity parameters instead of individual firms. Furthermore, the point of departure of distinct characteristics of exporters compared to non-exporters

[1]EuroGAP is the predecessor of GlobalGAP.

remains similar at the country-level.

3.2.1 Quantile regressions: estimation issues

As the key novelty of this analysis, quantile regressions are applied within the gravity framework as these allow us to distinguish heterogeneous effects of covariates at different parts of the distribution of the dependent variable. Thus, the empirical analysis is not constrained towards the conditional mean as it is the case with standard regressions.[2] We estimate the following exponential model:

$$Q_\tau \left[\ln Y_{ijpt} | \mathbf{X}_{ijpt}\right] = \exp\left(\mathbf{X}_{ijpt}\beta\left(\tau\right)\right), \qquad \tau \quad \epsilon \quad (0,1) \tag{3.1}$$

where Q_τ denotes the $\tau - th$ conditional quantile of $\ln Y_{ijpt}$. For example, for $\tau = 0.5$, $\beta(\tau = 0.5)$ is the coefficient at the median of the bilateral trade flow distribution. We estimate β for deciles.

Anderson *et al.* (2003) show that previous gravity estimations mostly do not account for third-country effects, consequently biasing the coefficients. The so-called "multilateral resistance" implies that bilateral trade between two countries is affected by the accessibility of both countries by third countries. Fixed effects are therefore frequently used to account for multilateral resistance (Head *et al.*, 2014). In our context, this implies the inclusion of at least importer-year, exporter-year, and country-pair-product fixed effects. Shingal *et al.* (2016)[3] argue that country-year-product fixed effects shall be included to address endogeneity at the sectoral level adequately. However, in this chapter, we consider importer-year, exporter-year as well as country-pair-product fixed effects which account country-year and for country-pair-product specific characteristics. Thus, we address multilateral-resistance equally well at the sectoral level.

Due to the size of our dataset, we do not estimate all fixed effects as additional regressors in Equation 3.1, but rather follow the procedure of Guimaraes *et al.* (2010) in which fixed effects are estimated beforehand. They are then subtracted from the dependent variable and subsequently, a quantile regression procedure is applied by using this transformed dependent variable. Although the procedure is computationally intensive and not efficient, it imposes minimum memory requirements.[4] This alternative procedure of computing fixed effects has been pointed out by Head *et al.* (2014) and applied in different fields as well; see for example Pomeranz (2015) for an analysis about taxation enforcement. McCaffrey *et al.* (2012) summarise different techniques and Stata commands that have been developed recently to control for unobserved heterogeneity. See Appendix B.1 for more methodological details.

[2] Quantile regressions are also beneficial from a methodological point of view as discussed in detail by (Figueiredo *et al.*, 2014)

[3] This article has not been included in this dissertation because our co-author Liliana Foletti has already included parts of it in her dissertation. It is currently under review at American Journal of Agricultural Economics.

[4] Therefore, we use the Stata command `reg3hdfe` which is a modified version of `reg2hdfe`. We are thankful for Prof. Figueiredo who provided the code of `reg3hdfe`.

The gravity model as mentioned above naturally assumes strictly positive trade values. However, most trade datasets are characterised by a large fraction of zero trade flows, especially for highly disaggregated data (Head *et al.*, 2014). Zeros emerge either due to statistical rounding, missing values, or because of *de-facto* zero trade. Statistical rounding is highly common in trade data and hence represents one major source of zeros. Consequently, we observe bilateral trade flows Y_{ij} between both countries as long as $\ln Y_{ij}$ exceeds the threshold κ:

$$Y_{ij}^{*} = \begin{cases} Y_{ij}, & \text{if} \quad \ln Y_{ij} > \kappa \\ 0, & \text{if} \quad \ln Y_{ij} \leq \kappa \end{cases} \tag{3.2}$$

As we describe in Section 3.2.2 below, our dataset exhibits 84% zero trade values. Because all zero-observations would be dropped when takings logs, which is a necessary transformation for the estimation of gravity models via OLS, this could lead to a systematic selection bias. Not accounting for this high frequency of zeros in the estimation would hence yield inconsistent estimates. Cameron *et al.* (2010) and Head *et al.* (2014) suggest to replace zeros by the minimum uncensored value of $\ln Y_{ij}$ to address this challenge adequately.

The remaining problem is that we do not know which observations are actually censored. Therefore, we follow a three-step procedure proposed by Figueiredo *et al.* (2014, pp.10-14).

I In a pre-step, we compute the fixed-effects $\mu_{ijp}^{1}, \eta_{it}^{1}, \nu_{jt}^{1}$ which we need to estimate the probability Π that two countries are engaged in trade by using standard gravity variables like GDP, distance, tariffs, language etc. with a simple linear probability model:

$$\Pi(\mu_{ijp}^{1}, \eta_{it}^{1}, \nu_{jt}^{1}, \mathbf{X}_{ijpt}) = \Pr\left(Y_{ijpt} > 0 | \mu_{ijp}^{1}, \eta_{it}^{1}, \nu_{jt}^{1}, \mathbf{X}_{ijpt}\right) \tag{3.3}$$

The estimated probability $\hat{\Pi}$ is then used to construct the subset J_0:

$$J_0 = \left\{(i,j,p,t) : \hat{\Pi}(\mu_{ijp}^{1}, \eta_{it}^{1}, \nu_{jt}^{1}, \mathbf{X}_{ijpt}) > 1 - \tau + c_N\right\}$$

II Again, we compute the fixed effects $\mu_{ijp}^{2}, \eta_{it}^{2}, \nu_{jt}^{2}$ beforehand. Since we do not estimate a linear probability model as in step one but rather a quantile regression model instead, $\ln Y_{ijpt}$ is the dependent variable. As proposed by Guimaraes *et al.* (2010) we use a transformed dependent variable which is the difference between $\ln Y_{ijpt}$ and the estimated fixed effects: $\ln \hat{Y}_{ijpt} = \ln Y_{ijpt} - \hat{\mu}_{ijp}^{2} - \hat{\eta}_{it}^{2} - \hat{\nu}_{jt}^{2}$. The fixed-effect quantile regression estimator of Koenker (2005) is applied to the observations of subset J_0.

III Finally, the third step basically repeats the second one to increase efficiency. We construct a second subset J_1 by using the quantile regression parameters $\hat{\beta}^{0}$ of step two:

$$J_1 = \left\{(i,j,p,t) : \hat{\mu}^1_{ijp} + \hat{\eta}^1_{it} + \hat{\nu}^1_{jt} + \mathbf{X}_{ijpt}\hat{\beta}^0) > 1 - \delta_{NT}\right\}$$

with $J_0 \subset J_1$ and where δ_{NT} is a small positive value. Step two is repeated by using observations of the subset J_1. Hence, fixed effects based on this subset are computed and subtracted from the dependent variable. This transformed variable $\hat{Y}'$ is then used as dependent variable in the final quantile regression.

To conclude, our estimate procedure accounts for all mentioned empirical challenges associated with the gravity model. First, we include country-year and country-pair-product estimates to address potential concerns related to multilateral resistance. Second, we employ a comprehensive three-step procedure to account for the selection bias. Third, quantile regressions naturally account for heteroskedasticity due to their property of being invariant to monotone transformations.[5]

The final specification of the empirical model is then given by Equation 3.4.

$$Q_\tau\left[\ln\hat{Y}'_{ijpt}|\mathbf{X}_{ijpt}, \mu_{ijp}, \eta_{it}, \nu_{jt}\right] = \beta(\tau)_0 + \beta(\tau)_1 \mathrm{R}_{jpt} + \beta(\tau)_2 \ln(\mathrm{Tariff}_{ijpt} + 1) \quad (3.4)$$

Note, that $\hat{Y}'$ is the transformed dependent variable which is a result of the three-step procedure. The main explanatory variables are the restrictiveness indices (R_{jpt}) as described below by Equation 3.5 and tariffs (Tariff_{ijpt}). The latter do vary along the importer-exporter-product-year dimension and are therefore, not captured by the fixed effects. We estimate Equation 3.4 based on the aggregate dataset.[6]

3.2.2 Data

Our target variable is food standards and we use MRLs which determine the thresholds of tolerated residues of for example pesticides in food imports. MRLs are an important public food standard and they are frequently used in empirical research. MRLs are adopted to control for aflatoxin contamination of the food supply. The toxicity of aflatoxin for human and animal health has been analysed and confirmed for many decades (Eaton *et al.*, 1994). Aflatoxins are naturally occurring in mycotoxins that are produced by various species of fungi in agricultural commodities. For example the fungi *Aspergillus flavus* appears in soil, wheat, and hay and especially grows at high temperatures and humid climates. Aflatoxins are toxic and belong to the group of most carcinogenic substances. For example, the lethal intake of the aflatoxin B1 lies between 1-10 mg/kg of the bodyweight of an adult.

Therefore, in order to facilitate comparability to previous studies on the effects of food standards on trade flows, we use MRLs as a measure of the restrictiveness in terms

[5] A mathematical derivation of this property is omitted here. Instead, we refer again to Figueiredo *et al.* (2014).

[6] Regressions at the disaggregated product-level often faced problems related to multicollinearity and convergence such that the few results obtained were not representative and are omitted here.

of food standards of the importing country. Data are obtained from the Agrobase-Logigram's Homologa database which covers ten years from 2005 to 2014. In total, we include 54 countries. Table B.1 displays the number of MRL-observations per country and year. They differ substantially across countries and years. For example, Ukraine set only 154 different MRLs in 2005, but 454 in 2014. The total number of MRLs almost doubled from about 15,000 in 2005 to about 29,000 in 2014. Figures B.1 and B.2 in the appendix display the number of regulated residues and the number of regulated products per country for the years 2005 and 2014, respectively. European countries, Germany and Austria in particular, were leading in terms of regulated residues in 2005. Ten years later, all EU countries are now subsumed as EU due to the harmonisation of MRLs in September 2008, regulated more than 500 residues per product on average. Japan regulated more MRLs than any other country in 2014. Figure B.2 also shows that the US regulated more than 900 products in 2005 followed by Australia, Germany, and other European countries. In 2014, Candada regulated most products while the US reduced its number slightly but still exceeds 800. Upper-middle-income countries such as Mexico, Brazil, South Africa, and Argentina regulate on average fewer products and residues.

Countries usually apply a mixture of country-specific MRLs, default values, and international harmonised values. The plain dataset only contains country-specific MRL data. Thus, additional adjustments are required to obtain the complete dataset. However, to make our results more robust to the way of dataset construction, we use three different datasets: RS0 indicates no replacements while RS1 and RS2 are constructed via a clear procedure which is explained below.

In RS0, missing observations were not replaced at all. In RS1, missings were replaced with country-specific values according to the following procedure, see Table B.2:[7] Some countries apply the Codex Alimentarius (FAO and WHO, 2016) as international harmonised MRLs if nothing else is specified. The Codex Alimentarius is an international harmonised food standard system which was developed by the Food and Agricultural Organization (FAO) and World Health Organization (WHO). Hence, for specific countries we replace missing observations with values of the Codex Alimentarius. Other countries, for example Canada, apply the default value of 0.01 ppm if nothing else is specified. Furthermore, we replaced missing values of EU members states with 0.01 parts per million (ppm) for the years 2009 to 2014 since the EU harmonised the MRLs from September 2008 onward. We have country-specific data for most EU member states and data for the EU overall. However, some EU member states such as Latvia, Croatia, Finland, or Slovenia were not included in the purchased dataset. Thus, we use EU values for these countries. Finally, we replace missing values of each importer-product-year-residue combination with the maximum value (i.e. the least restrictive value) across all importing countries to create the third dataset RS2. This procedure is most frequently used in articles that construct restrictiveness indices, for example Ferro *et al.* (2015). Therefore, RS2 is our main dataset in the following sections and analyses based on RS0 and RS1 were only employed as robustness-checks.

[7] This replacement-procedure is based on additional information that we received from Homologa.

Restrictiveness index

A direct inclusion of MRLs in the regression analysis would lead to misleading results. We are interested in the *de facto* restrictiveness of MRLs for a specific product p in time t of the importing country j. However, an absolute value for example of 0.5 ppm for pesticide A might be less or more restrictive in practice than for pesticide B. Furthermore, we need to aggregate across residues to make the computation feasible. For example in the case of MRLs of Spain in 2014, we have data for 779 different residues just for Maize. Thus, we need an index that aggregates across residues and normalises MRLs.

Ferro *et al.* (2015) developed a restrictiveness index which is normalised between zero and one and defined for the importing country j, product p in time t as:

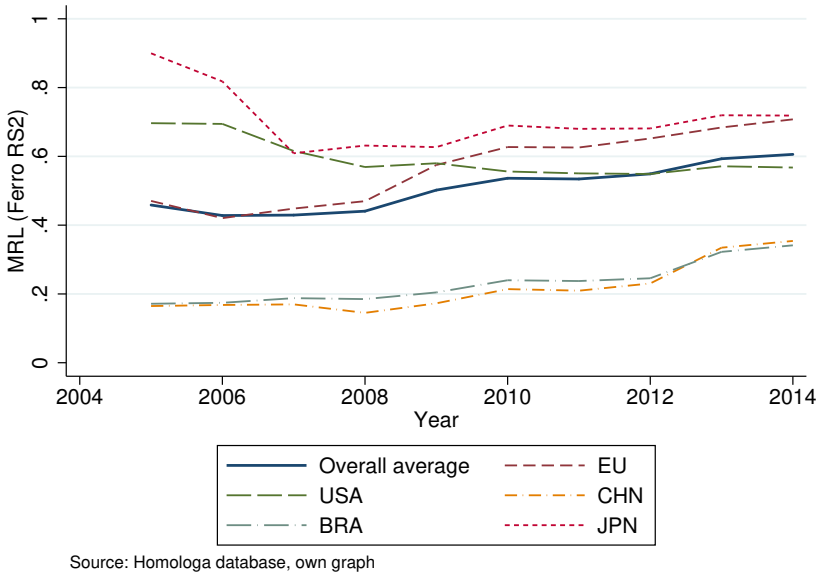
$$\text{Ferro}_{jpt} = \frac{1}{N(a)} \sum_{n(a)=1}^{N(a)} \frac{\text{MAX}_{p,a,t} - \text{MRL}_{j,p,a,t}}{\text{MAX}_{p,a,t} - \text{MIN}_{p,a,t}} \tag{3.5}$$

$\text{MAX}_{p,a,t}$ refers to the maximum MRL of product p, pesticide a in year t across all importing countries. Values closer to unity imply stricter standards and hence, the index does not require an inverse interpretation as MRLs in general do. The drawback is that the economic interpretation becomes less accurate by using this index. It limits the interpretation towards sign and statistical significance of the coefficients. Also, while we can compare magnitudes across coefficients and to coefficients of other studies, we will not know to what extent a change of MRLs relates to expected changes of bilateral trade flows.

The development of Ferro indices are depicted in Figure 3.1. The solid line represents the average across all countries and products whereas the other lines represent averages across products of different countries (or country groups). There is a slight increasing trend of the overall average which implies that MRLs are set at relative stricter levels over time. The increase is in particular pronounced from 2008 to 2009 which is partly driven by the harmonisation of MRLs of the EU (red long dashed line). The EU set continuously stricter MRLs and overtook the US in 2009 which became slightly less strict over time. China, which has been significantly less strict on average than other countries, has been catching up in 2013 and 2014. Although Japan's restrictiveness has declined remarkably from 2005 until 2007, it has remained among the strictest countries.

Developed countries are usually criticised for imposing relatively strict standards on their agricultural imports. Figures 3.2 and 3.3 provide general support for this view by relating number of regulated MRLs and restrictiveness to per capita income. Figure 3.2 plots the average number of residues of a country and per capita income calculated as the mean from 2005 to 2014. Richer countries impose more MRLs in quantitative terms than poorer countries. Moreover, richer countries impose not only more MRLs but also relative stricter standards. The two strictest countries, however, are outliers. Vietnam and South Africa set the strictest MRLs on average. BRICS-countries (except South Africa) are less restrictive in relative terms: Brazil, Russia,

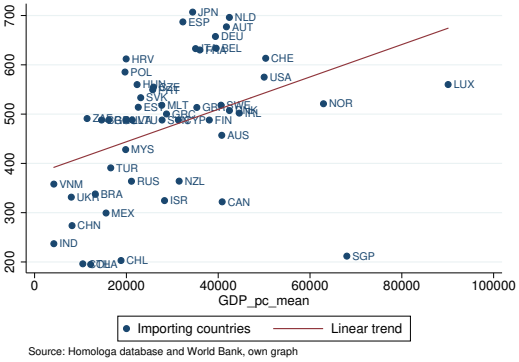
FIGURE 3.1: The restrictiveness index based on Ferro *et al.* (2015) over time for selected countries.



Source: Homologa database, own graph

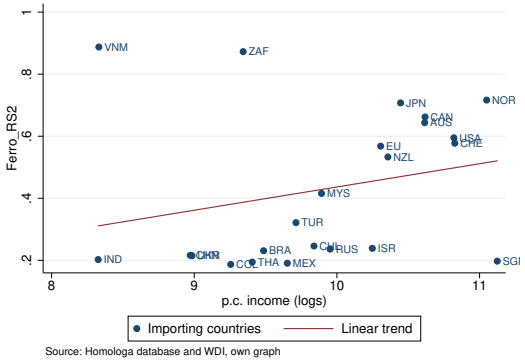
India, and China are located in the bottom.

FIGURE 3.2: Number of regulated residues and p.c. income



Source: Homologa database and World Bank, own graph

FIGURE 3.3: Restrictiveness (Ferro, RS2) and p.c. income



Source: Homologa database and WDI, own graph

Figures B.3 and B.4 in the appendix show worldmaps in which dark shaded countries either regulate many residues or set strict standards. The EU as well as the USA set many MRLs as well as relatively strict standards. Overall, Canada, Argentina, Norway, and Japan set the strictest MRLs whereas most other Asian countries are less strict. Data for African countries and most countries located in the middle-east are rare such that these regions are (except for Egypt and South Africa) are excluded from the analysis.

Since the analysis is conducted within the gravity framework, the remaining variables are standard in gravity models. We use import data for 54 importing and exporting countries for which we have MRL data in current US-Dollars from the UN Comtrade database. Overall, we selected 49 products depending on the number of

TABLE 3.2: List of included products

HS Code	Product	HS Code	Product	HS Code	Product	HS Code	Product
080211/2	Almonds	0707	Cucumbers	100820	Millet	1007	Sorghum
080810	Apples	070320	Garlic	1004	Oats	1201	Soybeans
080910	Apricots	080540	Grapefruits	071120	Olives	070970	Spinach
070920	Asparagus	0806	Grapes	070310	Onions	081010	Strawberries
070930	Aubergine	080221/2	Hazelnuts	080510	Oranges	0902	Tea
080440	Avocados	1210	Hop	080720	Papayas	24	Tobacco
0803	Bananas	081050	Kiwi	080930	Peaches	0702	Tomatoes
1003	Barley	080530	Lemons	080820	Pears	080231/2	Walnuts
070410	Broccoli	0705	Lettuce	080430	Pineapples	080711	Watermelons
0706	Carrots	1005	Maize	080940	Plums	1002	Rye
070940	Celery	080520	Mandarins	0701	Potatoes	1001	Wheat
080920	Cherries	080450	Mangos	081020	Raspberries		
080240	Chestnuts	080710	Melons	1006	Rice		

MRL-observations. Table 3.2 provides a list of all products included and the corresponding HS codes.

Ad valorem equivalent tariffs are obtained from the International Trade Center[8] for each trading pair, product, and year. Other typical gravity variables like proxies for the masses of both trading partners and proxies for trade costs such as distance, common language, shared colonial history, contiguity, and religion are not included because these vary only at the country-pair dimension and are captured by country-pair-product fixed effects.

3.3 Main results: The effect of MRLs on trade

This section reports the main results of the quantile regression of Equation 3.4. Figure 3.4 shows coefficients of restrictiveness (solid line) across deciles based on the index of Ferro *et al.* (2015) and the RS2 full dataset. The full set of results are presented in Table B.3. The coefficients of MRLs remain below zero for all deciles except for the ninth decile. The coefficient for $\tau = 0.1$ is close to minus two, increases to -0.13 in the third decile, declines again to more than -1.5 in the median regression and increases then steadily until 0.5 in the ninth decile. Moreover, the coefficient is statistically significant for all deciles.

Although the index does not allow us to infer on economic significance directly, we can compute trade changes as deviations from the mean. The mean of the index is equal to 0.51. If the index increased from the mean by 0.01 to 0.52, trade between country-pairs at the lowest decile declines by a factor of $e^{-1.973 \times 0.01} = 0.98$ which implies a decline in trade flows of about two per cent. In contrast, trade flows between country-pairs at the largest deciles are expected to increase by 0.5 per cent if the mean increases by 0.01.

This supports the trade-barriers perspective for most trading country-pairs except the highest decile. Hence, MRLs impede bilateral trade between country pairs with

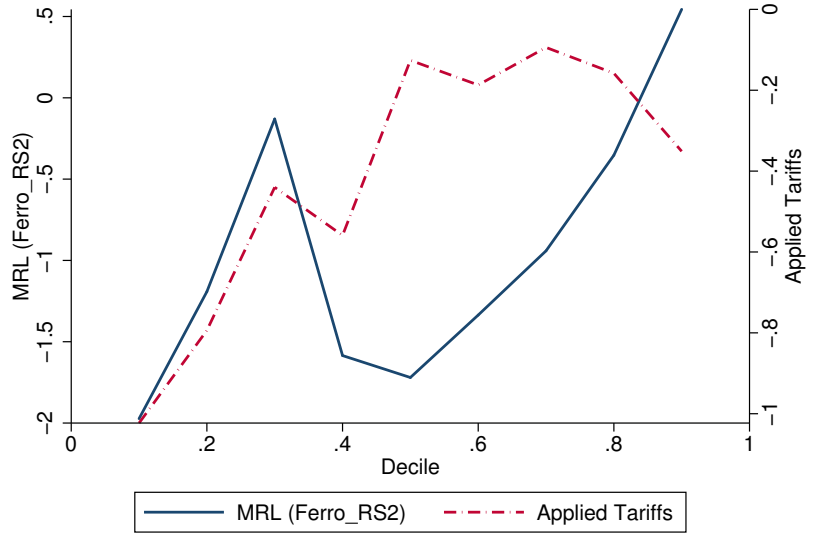
[8] Access via www.macmap.org in January 2016.

relatively low trade volumes including the pair at the median. However, the catalyst-to-trade view predominates for the highest decile. Thus, the effect of MRLs on trade is heterogeneous which we interpret as empirical support for our hypothesis and justifies quantile regression as our preferred method.

The coefficients on the variable tariffs are depicted by the red dashed line. Again, we find heterogeneous effects which underlines the importance of applying quantile regressions in this context. Tariffs are found to reduce trade independent of the respective decile. The coefficients are statistically significant for all deciles. However, the trend is less clear compared to MRLs because the effect becomes again more negative after a sharp increase in the eighth and the ninth decile.

Results that are based on the other two datasets RS0 (no replacement of MRLs) and RS1 (replacements with country-specific MRLs only) are less clear (Table B.3). The coefficients of the lowest decile remain of similar magnitude of about minus two. The coefficients of the highest decile are larger than of the lowest decile for all datasets but only positive for RS2. The effect becomes less negative from the fifth to the eighth decile in all three datasets. In contrast to RS2, country-pairs are found to be more adversely affected by standards at the ninth decile than for the eighth decile as long as RS0 and RS1 are considered. Overall, we do find heterogeneity of the effect of food standards on trade, but the trend is less clear for the other datasets RS0 and RS1. Overall, we expect RS2 to be closest to the true degree of restrictiveness because it is reasonable to assume that countries, that do not have a regulation of a specific residue, are likely to be at least as restrictive as the least restrictive country. Moreover, not replacing any MRL (RS0) is unrealistic as well because it ignores valuable information on default values. These findings underline the sensitivity of results of the empirical literature of MRLs and trade.

FIGURE 3.4: Estimated coefficients of restrictiveness (left axes) and tariffs (right axes) per decile of the trade flow distribution.



Source: Agrobase–Logifram Homologa database and ITC
Authors' estimation

3.4 Discussion and Conclusion

Based on methods reflecting the current research frontier in empirical trade analysis, we find evidence that the effect of food standards on trade flows differs by export size. We find a similar, heterogeneous effect for tariffs, a more traditional trade cost variable. We base our argument on two pillars: First, food export markets are highly concentrated. Few countries often account for more than 50% of total world exports at the product-level. We show that the combination of concentrated agricultural export markets and compliance costs due to food standards lead to heterogeneous effects of standards on trade.

Standard regression techniques, which do not allow for differentiate effect sizes, would have missed the heterogeneity of the effect of food standards on trade flows conditional on the size of exports. In line with previous studies, we would have been able to estimate the average effect only, which turns out to be negative in this case. However, quantile regressions allow us to draw more specific conclusions: It is not only income that matters for compliance but also export volume.

In terms of policy recommendations, we emphasise that provision of fair and equal market access is certainly necessary, but not sufficient to allow all countries to benefit from globally integrated agricultural export markets. Aid-for-trade programmes are one option which especially help those countries which have the potential to enter global export markets but are constrained by the high fixed costs of compliance. Nevertheless, since we find that the *ex-ante* export volume – among other factors – determines whether or not countries comply with stringent food standards, not all countries will gain from global export markets in the same way.

The results also help to improve our understanding of how food standards affect the structure of global agricultural markets. Our findings touch upon results of the current Agricultural Outlook 2015-2024 of the Organisation for Economic Cooperation and Development (OECD) and FAO in which the high degree of export concentration is identified as a serious risk (*Agricultural Outlook 2015-2024*). As specialisation processes of food production continue, food importing countries become more dependent on a very small set of exporting countries. If countries with smaller trade volumes continue to be more adversely affected by stricter standards, the number of exporting countries per product might decline even further. In the long term, global agricultural export markets might become even more concentrated if standards continue to become stricter. One could even think of a vicious cycle in the sense that more concentrated agricultural export markets are, the larger the negative effect is for country-pairs with smaller trade volumes. This might be particularly relevant if fixed costs of compliance are not only of a financial nature. For example, the more experienced countries are and the larger their export volumes, the more product-specific knowledge they could accumulate over time with corresponding spill-over effects which, in turn, increase the comparative advantage of those few countries even further.

If agricultural export markets continue to become more concentrated, worldwide supply of particular products might become more vulnerable to external shocks. If, for example, extreme weather conditions in California reduce the harvest of almonds,

the effect on worldwide supply of almonds would be substantial because only very few other producing countries would be able to replace the USA as the major exporting country. Although the consumption of almonds might be less essential than other products, as for example onions in India, it illustrates possible drawbacks when agricultural markets become increasingly concentrated. A shortage of supply of politically sensitive products is likely to result in protectionist trade policies like the recent export ban of Indian onions (Bhosale, 2017). Increasing volatility of supply could increase food price volatility with direct effects on poverty of affected countries. From a market-competition perspective, a sufficient set of exporting countries is desirable to allocate production risks adequately. High degrees of concentration increases the likelihood that a small set of suppliers can ask for an extra price-premium which contradicts gains of free and liberal markets.

Although the pattern of heterogeneity of the effect of tariffs on trade is similar compared to that of MRLs, it requires a different discussion. The essential argument for finding the heterogeneous effect of MRLs on trade builds upon fixed costs of compliance. In contrast, ad valorem tariffs are not characterised by fixed costs but rather depend on total trade value. Moreover, there are no obvious arguments for trade-enhancing effects of tariffs on trade as it is the case for standards. Hence, our finding of purely negative effects of tariffs on trade matches our hypothesis that fixed costs of compliance make the crucial difference.

However, the high degree of market concentration as the second argument which underlies our hypothesis of heterogeneous effects remains valid for tariffs as well. Country-pairs with low trade volumes are more vulnerable to tariff increases, probably because these trade relationships are less well established. On the other hand, country-pairs at the upper part of the trade flow distribution with well established trade relationships might be less affected as tariffs increase because they trade regardless. In other words, their trade elasticity with respect to tariffs is lower.

If this explanation is true, we should find similar patterns of heterogeneity of other trade cost proxies as well. We find some evidence in Table 4 of Figueiredo *et al.* (2014) in which the authors analyse the effect of WTO membership on trade. They include a dummy for common currency in their analysis, which reflects trade costs due to different currencies of trading partners. The coefficients for the lowest quantile and the median regression are both equal to 0.8, but increase to one for the upper quantile. However, the discussion of heterogeneous effects of all types of trade cost proxies is left to further research as it exceeds the scope of this paper.

Policy recommendations are not limited to exporting countries. The overall gain from setting standards can be expected to diminish. The SPS and Technical Barriers to Trade (TBT) regulatory framework of the WTO already provides an adequate mechanism to ensure that public standards are set based on scientific evidence. Allowing each country to set their individual food safety standards makes trade costly. Moreover, we can show that high-income countries tend to set stricter MRLs compared to lower-income countries. Thus, high-income countries impose an additional burden on international trade which is in particular problematic for countries with lower trade volumes. As shown in earlier studies and discussed in Section 1.2, historical trade

flows affect current trade flows. However, their effect seems to be relevant in particular once trade costs, especially fixed costs due to compliance with standards, increase. The effect of past trade flows on current trade is not desirable from a liberal market perspective if it impedes entry to global export markets. Other arguably aspects such as individual productivity levels or the ability to address relevant preferences should determine entry decisions of countries and firms and shall not be overweighted by past trade volumes. This is a matter of concern even if countries with low exports of a particular product are not necessarily low-income countries: as we will show in Chapter 5, the level of income of the exporting country is relevant as well for the effect of standards on trade.

Therefore, standards can be problematic from a development perspective and potentially counteract aid-for-trade programmes which aim to integrate developing countries into global value chains. Although Swinnen *et al.* (2015) argue that standards have the potential to facilitate trade integration processes, we show that stricter MRLs impede integration strategies based on proliferation of certification schemes as described by Schuster *et al.* (2013).

Countries which set strict MRLs should take the described negative consequences for other potential trading partners into account. Setting the "right" MRL shall be the objective and not setting MRLs at levels that are too strict on average. The determination of the correct MRL requires solid and long-term research of related disciplines. Policy makers should guarantee that sufficient research can be conducted in adequate environments.

Costs for compliance, but also for determining the correct MRLs can be minimised by harmonisation of standards across countries. Thus, the harmonisation of MRLs within the EU in 2008 was a useful step into this direction. Instead of allowing each country to set its individual MRL, the WTO might set incentives as such that member countries apply the Codex Alimentarius (or a modified version) more frequently. There is no a-priori argument based on food safety concerns why MRLs should differ across countries. Food-safety requirements can be assumed to be identical – or very similar at least – for every human being. Nevertheless, preferences differ across countries which makes complete harmonising of standards difficult and not desirable. In addition, because harmonisation could either lead to stricter standards or to increasing protests of citizens in countries that end up with less strict standards. These concerns would certainly need to be taken into account and not neglected.

To conclude, our findings explain at least partially the contradicting results in the literature whether standards act as catalysts or barriers to trade. We contribute to this debate by providing empirical evidence for the presence of *both* effects. Future research should address which conditions are required to support the catalyst-effect: How shall standards be designed to be least-distorting for trade? That is, how can risk of trade-distorting and trade-reducing effects be minimised for countries that trade little? Furthermore, a broader set of standards needs to be examined to improve external validity of results. For example, standards that indeed enhance trade at the conditional mean, could potentially also affect trade heterogeneously.

4 Institutional quality and the ability of organisations to adopt foreign agri-food trade standards

Abstract:

The implementation of quality standards among agricultural exporters is often described as having fixed cost character. This can be misleading if fixed costs are only understood in terms of required monetary investments. Instead, standard adoption can be viewed as the result of exporting countries' private and public organisations managing to solve the standard implementation problem. This hypothesis is tested based on the Worldwide Governance Indicators as proxies for the institutional characteristics of countries that successfully export fruits, dairy products, meat, fish, and vegetables. Within a gravity framework, the role of export countries' institutional quality is assessed for agri-food exports to the EU as a market with relatively high standards, in comparison to their agri-food exports to all markets. Results indicate that institutional quality is not relevant for successful exports of fruits, vegetables, and fish, while it turned out to be an important determinant for successful exports of dairy- and meat products, which may be viewed as products of rather demanding process- and quality standards.

This chapter is joint work with Sebastian Hess. An earlier version has been published in: Perspektiven für die Agrar- und Ernährungswirtschaft nach der Liberalisierung, Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaus e.V., Band 51, 157-170, 2016, *Landwirtschaftsverlag Münster*.

Malte Ehrich's contributions are: Development of research question, theoretical and empirical framework, data management, as well as empirical estimation. Both authors contributed to the research design and both wrote the manuscript.

Acknowledgements: We are thankful for valuable comments of Thomas Hertel, Thomas Heckelei, Roland Herrmann, Jens-Peter Loy, and Bernhard Brümmer.

4.1 Introduction

Fixed costs of compliance with food standards remain the major reason why food standards might be a threat for global agricultural export markets (see Section 1.2). However, the nature of fixed costs is poorly understood. Within the international trade policy environment, distinctions between different standards and their varying influence on fixed costs of trade are not commonly considered. Especially exporting developing countries tend to view food quality standards broadly as nontariff barriers that impede their market access (WTO, 2013).

At first glance, the concept of fixed cost suggests that financial investment in some way or the other (e.g. aid for trade programmes, foreign direct investment in the exporting country) would be the main constraint to overcome such trade impediments. In this context, the WTO (2013) conducted a private sector survey in which they evaluated several aid-for-trade programmes. The authors identify various non-financial barriers of standard adoption, e.g. the absence of networks with governments, the absence of supply chain relationships, a lack of confidence in existing and new markets, and difficulties in collaborating with the public sector (WTO, 2013, p.13). Clearly, the provision of financial support alone would not be sufficient nor adequate to eliminate these non-financial trade barriers.

Nunn *et al.* (2013) and Levchenko (2007) identify the general quality of public institutions as a potential source of comparative advantage. Nunn (2007) analyses the role of contract enforcement and its implications for product-specific investments and finds that institutional quality alone explains an even larger share of trade flows than skilled labour and capital combined. Institutional aspects of trade have been described in terms of contract enforceability, democracy, social networks, relevance of informal activities, information asymmetries, and the availability of skilled workers e.g. by Rauch *et al.* (2003), Acemoglu *et al.* (2007), Yu (2010), Bloom *et al.* (2012), de la Mata, T. *et al.* (2013), Francois *et al.* (2013), and Muange *et al.* (2014).

Henson *et al.* (2004) underline the relevance of public institutions in the case of Kenyan Nile perch exports. The authors identify weak governance as one reason why only few Kenyan exporters upgraded their production facilities as hygiene standards of the EU between 1997 and 2000 became stricter. Although the Kenyan government "invested heavily in changes to legislation, administrative structures, [...], and certification procedures", other factors such as infrastructure remain of low quality and increased fixed costs of compliance by about 25% (Henson *et al.*, 2004, p.75).

Thus, well-functioning public institutions seem to increase trade in general, while the "fixed cost" behaviour of food standards seems to be more complex, and potentially related to different institutional aspects of the firm or the country that aims to adopt a certain standard. Alternatively, different types of food standards or standards for different types of food products might reveal different degrees of sensitivity with respect to their trade-elasticity of institutional quality.

We therefore study the particular effect of institutional quality on developing countries' probability to successfully export agri-food products to high-value markets by

analysing the product-specific effect of the exporting country's quality of public institutions.

It is hypothesised that the effect of institutions on exports depends on the one hand on the stringency of food standards of the importing country and on the other hand on the degree of complexity of the exported product. If this hypothesis was correct, then high quality institutions in an exporting country should *c.p.* prove relevant for agri-food exports to markets with relatively strict standards, and for agri-food products with relatively demanding standards.

This article is therefore organised as follows: the next section explains our conceptual framework from which we derive the hypotheses to be tested. Section 4.3 presents data and the different gravity model specifications that are used. Results are presented in section 4.4. Section 4.5 discusses the results and concludes the chapter.

4.2 Conceptual framework

The models by Caliendo *et al.* (2012) and Garicano *et al.* (2015) focus on the role of firms' ability to organise knowledge-intensive production optimally in order to benefit from export activities. Even though these models do not consider the problem of standard implementation explicitly, they can provide a theoretical foundation for the relationship between institutions and export success: in both models, production requires knowledge (as an interpretation of capital) and labour (to be thought of as the amount of time that workers are willing to offer to the labour market). One unit of labour (e.g. 1 hour) can be supplied by two types of agents: workers (L) and managers (K). Both together solve problems for which a minimum level of knowledge is required. Problems need to be solved in order to transform otherwise useless production into α, which is useable output of the corresponding product variety.

We therefore assume that in a given country, the agri-food sector is composed of heterogeneous firms. A firm pays fixed costs f^{α} in order to design a product variety α. Melitz (2003) argues in this context that only the most productive firms will export. However, instead of randomly drawing a productivity parameter as in the model of Melitz (2003), here the corresponding total market demand for product variety α is by each firm drawn from a probability distribution. This level of demand induces the optimal organisational structure in terms of optimal share of workers L and managers K in labour that the firm has to match. In other words: Higher (global) market demand for a product variety α will require firms to work with a higher level of problem solving capacity than firms with lower market demand for their variety.

Garicano *et al.* (2015) show this in their theoretical model, where the purpose of a hierarchical organisation (such as e.g. an exporting firm) is to economise on the acquisition of knowledge in order to solve problems, and this is obviously more relevant if the costs of acquiring relevant knowledge are relatively high. Thus, firms that are better at solving problems will be more productive, and consequently are going to be more likely involved in export activities.

For instance, in order to solve the problem of implementing a new standard relevant for entering a foreign market, the firm will have to invest effort in terms of capital, time, and knowledge. In other words: implementing the new standard will increase the demand for managers K in the firm relative to workers L, and the fixed costs related to the export of α will raise consequently, which on turn can pay off only if a comparatively large market is served, namely through exports.

We therefore conclude that the fixed cost character of a foreign trade standard can be distinguished according to a monetary component and an organisational component. This organisational component should consist of a private and a public dimension: the private dimension reflects (change in) organisational capacity within the exporting firms, e.g. more managers are hired who have better knowledge and skills about how to implement the new standard.

The private dimension of export-related organisational capacity within any firm should according to Garicano *et al.* (2015) be reflected by the firm's productivity, and should therefore correspond to the market demand for α. This private dimension of organisational capacity should under conditions of monopolistic competition or perfect competition thus in the long run just be equivalent to the firm's productivity. Since this is primarily the outcome of competition between firms, we are assuming that this private dimension of organisational capacity remains constant and does not get affected by the public dimension. Instead, we are especially interested in this public dimension of export-related organisational capacity.

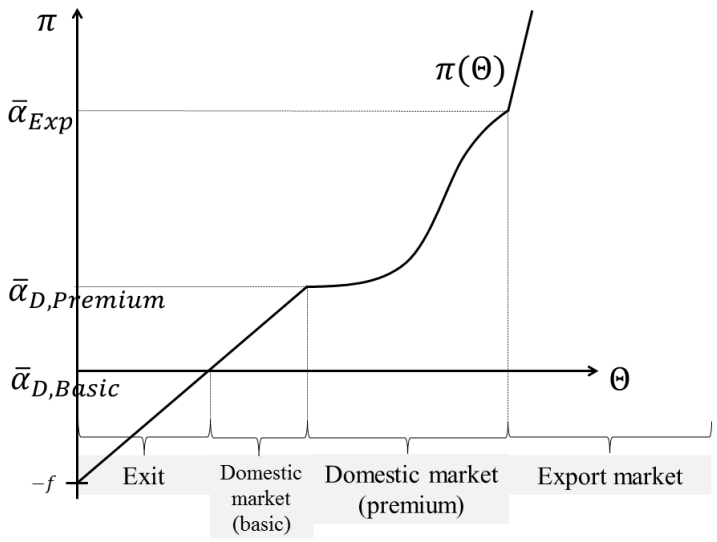
The public dimension however describes the institutional environment within which the firm's managers have to solve the problem of implementing the new standard for α. Obviously, two identical firms within two different institutional environments will perform differently: institutions that are relevant for exporting activities are a country's foreign missions, the diplomatic body, access to credit and insurance markets, quality of communication infrastructure, freedom and transparency regarding legal procedures, and the level of corruption in the country where the standard has to be implemented.

This institutional component is in our analysis denoted by Θ and describes the "trade standard adoption ability' (SAA) of a firm to successfully implement and adopt a certain quality standard for its exports to the standards-setting importer j. Based on the empirical literature about institutional aspects of trade, SAA (Θ) can be expressed as a function of various indicators of institutional performance:

$$\Theta = f(social\ capital, government\ effectiveness, corruption, \dots) \tag{4.1}$$

Figure 4.1 displays the positive relationship between profits $\pi(\alpha, \theta)$ of a firm that is producing product variety α, and its corresponding SAA, Θ. Larger values of Θ are required to export to markets with relatively strict standards. The corresponding thresholds are indicated as $\bar{\alpha}$.

FIGURE 4.1: Self-selection of firms into markets that vary in stringency of product-related standards.



Thus, to match a relatively high standard, SAA has to be higher than for the adoption of a relatively low standard, whereas "high" and "low" refer to the required problem-solving capacity. This degree of complexity of the problem will depend on i) the type of standard that is set by the importing country and ii) the ease at which this standard can be implemented into the production process of $\bar{\alpha}$.

This complexity of the standard implementation problem that an exporting firm faces has at least two important dimensions: restrictions on content of the imported food item, and restrictions on related production procedures and processes. The restrictions on content constitute e.g. MRLs or minimum requirements in terms of quality content. Restrictions on procedures may refer to cooling and storing regulations, traceability, special treatment of animals, and so on.

In reality, those importing countries with more demanding standards regarding food quality will restrict imports based on a combination of these two dimensions, which leads to the following hypotheses regarding exporting firms' SAA:

Hypothesis 1: Higher SAA favours exports to markets with relatively demanding standards; an exporting country is *c.p.* more likely to be observed exporting to destinations with relatively high trade-related standards if the SAA is higher than in other countries.

Hypothesis 2: Higher SAA favours exporting those agri-food products for which standard implementation is relatively more demanding than for other agri-food products. Agri-food products may not all require the same level of SAA, e.g. because different levels of processing or e.g. cooling requirements or specific residues.

Both hypotheses are in the following sections put to an empirical test.

4.3 Data and model specification

Both hypotheses could be tested by conducting surveys and compiling firm level data in order to assess the values of Θ empirically, e.g. from subjective estimates that workers and managers provide about the relative ease at which they deal with various foreign agri-food standards. However, firm level data across countries are difficult to obtain and managers are usually reluctant to reveal data about productivity, export success, and internal restructuring.

Instead, official statistics about trade flows and trade-related standards contain implicitly all relevant information about the ability of the organisations within a certain country to adopt foreign trade standards.

This enables us to test the hypotheses within a gravity framework, after controlling for the effect of distance, GDP, tariffs, and institutional quality as an approximation to the public dimension of SAA. For this, the following testing strategy is employed:

I If institutional quality matters for exporting firms' SAA, then institutional indicators should reveal positive and significant estimation coefficients regarding observed trade flows from exporters towards those markets that exhibit more demanding agri-food trade standards than the average country.

II Ranking countries precisely according to the ease, at which their agri-food trade standards can be matched by trading partners, is difficult. Instead, we assume that the EU can serve as an example of a market that has more demanding agri-food standards than the average country. We do not claim that the EU has the highest standards of all countries. However, the EU may serve as an example of a market that is both relatively demanding in terms of standards and due to its size and purchasing power at the same time reasonable attractive for most global exporters.

III Both hypotheses are tested by evaluating the estimation coefficients of empirical proxies for institutional quality with respect to their effect on exports to the EU, in comparison to exports to all countries for various products.

4.3.1 Empirical approximations to institutional quality

A broad and heterogeneous array of proxies for a country's institutional quality exists in this respect. For example, the World Bank provides doing-business indicators which are based on firm-level questionnaires. Among others, these indicators assess the number of documents required to start a business.

In addition, the Worldwide Governance Indicators (WGI) of the World Bank provide in this context an example of several indicators that may approximate the quality of a country's public institutions (Kaufmann *et al.*, 2010). Such indicators include stability and effectiveness of the government, the extent of corruption, public violence, and — among others — freedom and democracy. Table 4.1 provides an overview of all six indicators that are potentially relevant for the abovementioned hypotheses.

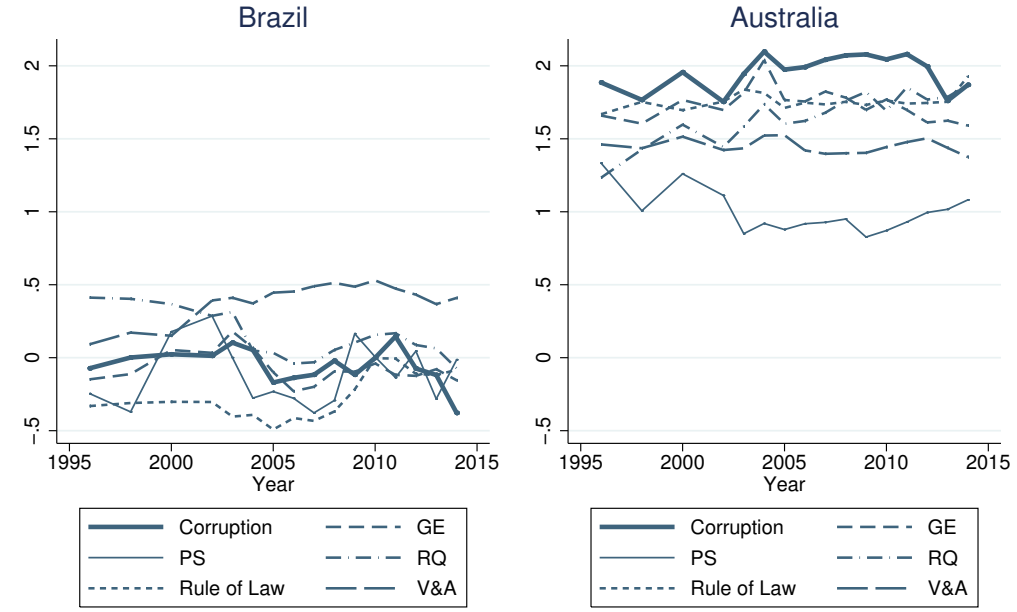
TABLE 4.1: Definition of WGI, Source: Kaufmann *et al.* (2010)

Indicator	Brief definition
(1) Control of corruption	Extent to which public power is exercised for private gain
(2) Government effectiveness	Perceptions of the quality of public services
(3) Political stability	Likelihood that the government will be destabilized or overthrown
(4) Regulatory quality	Ability of the government to formulate and implement sound policies
(5) Rule of law	Confidence in enforcement of contracts, property rights
(6) Voice and accountability	Participation in government, freedom of expression, association

For instance, higher government effectiveness at the country-level increases the SAA of firms because firms might receive more support for upgrading their production processes and for meeting foreign trade standards. In addition, the relevance of government effectiveness is even more pronounced for exports of relatively complex products for which standards are more difficult to fulfill.

Figure 4.2 plots all six indicators for Brazil and Australia, respectively. Both countries are major agricultural exporters. The WGI are heterogeneous between countries and vary over time. In this case, Australia achieves substantially higher levels of institutional quality such that we expect Australia to have a comparative advantage for exports of process-demanding products to the EU in particular.

FIGURE 4.2: Institutional quality varies over time (WGI range between -2.5 and 2.5), Source: Kaufmann *et al.* (2010).



GE: Government Effectiveness
PS: Political Stability
RQ: Regulatory Quality
V&A: Voice and Accountability

In addition, it turns out that all six indicators are highly correlated (Table C.1). Hence, including all indicators simultaneously into a gravity regression is not appropriate due to collinearity. Therefore, a principal component analysis (PCA) is applied instead of choosing only one indicator or an arbitrary subset. PCA allows us

to reduce the dimensions of the data. The PCA reveals that one component explains around 87% of the variation, see Table C.2. Corresponding eigenvalues are listed in Table C.3. This component represents institutional quality in the following model description and is denoted as WGI hereafter.

4.3.2 The Gravity dataset

The gravity dataset covers the period from 1996 until 2014. However, WGI are not available for 1997, 1999, and 2001. Bilateral imports in current US-Dollar are included from the UN Comtrade database of 170 importing and 177 exporting countries. In general, the quality of import data is higher than for export data because developing countries in particular rely on tariff revenues which are based on imports.

Five different categories of frequently traded agri-food products at different levels of processing were included into the dataset: Meat (HS 02), fish (HS 03), dairy products (HS 04), vegetables (HS 07), and fruits (HS 08). These products constitute categories of some of the most widely traded and at the same time highly valuable agri-food products. Furthermore, these products include sufficiently many different subcategories such that the implementation of different trade related import standards may constitute relevant organisational problems for the exporting firms. To keep the computation feasible, all country-pair-product combinations are excluded, for which less than three out of 16 observations over time are available which reduces the share of zeros from 85% to 36%. In total, the dataset contains n = 245,632 observations.

The standard gravity type variables have been obtained from the following sources: GDP of both trading countries are taking from the World Bank development indicators in current US-Dollars. Trade cost proxies like distance, colonial ties, common language, common border, and religion are downloaded from the CEPII homepage. In addition, ad valorem equivalent tariffs for each country-pair-product-year combination were taken from the International Trade Centre and have also been added to the gravity regressions. Table C.4 in the appendix presents descriptive statistics of the gravity data set.

4.3.3 The EU as an example of a relatively demanding import market for agri-food products

For all agri-food products on aggregate, the EU belongs to the world's leading agricultural exporter and importer regions. The EU's agri-food trade with non-EU countries has in principle to comply with all internal EU agri-food quality standards. The EU's general approach to food safety includes risk assessments, animal health, animal welfare, and plant health assessments that focus for many products on the entire supply chain "from farm to table".

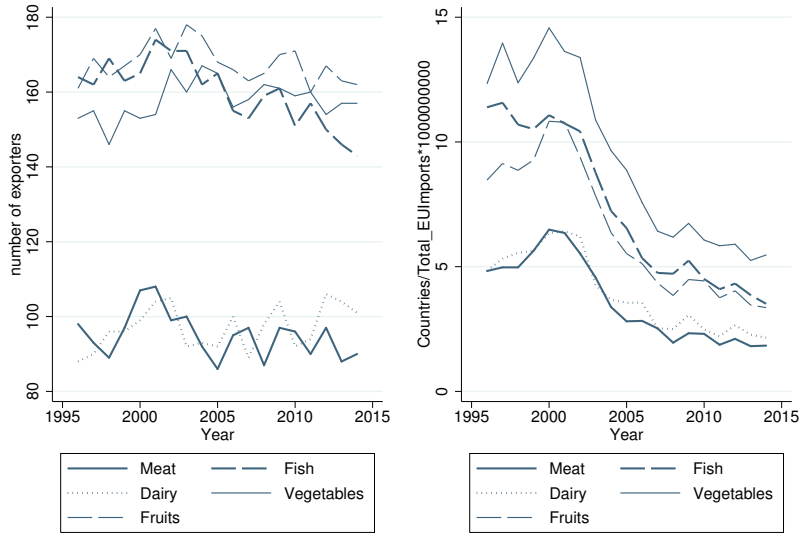
The European Commission for Health and Food Safety approves in this respect agri-food firms and processing plants in non-EU and EU countries (European Commission, 2017b). Thus, not only tariffs and non-tariff barriers limit the number of countries that have market access to the EU, but the cost and difficulties of implementing EU

standards and continuously having to comply with them considers a market entry barrier by itself.

The left panel in Figure 4.3 presents in this respect the number of observed trading partners that export in any year to the EU within each of the agri-food product categories under consideration. The left panel of this figure depicts the absolute number of exporting countries. The left-hand panel in Figure 4.3 reveals that only a relatively small number of non-EU countries is eligible for exports of meat and dairy products to the EU. In contrast, almost half of the world's countries are currently eligible for exports of fishery products, fruits, and vegetables.

The right panel shows these numbers relative to the total volume of imports of the EU in each of these product categories. This is especially meaningful because the EU external trade can be expected to have risen along both intensive and extensive margin (Hummels *et al.*, 2005). In addition, several EU enlargement steps took place during the observed time period, and further microeconomic shocks such as the creation of the euro zone or the financial crisis may have affected external trade flows.

FIGURE 4.3: Observed number of countries that export to the EU by product category, in absolute terms (left panel) and relative to total EU imports in these five product categories (right panel).



Source: UN Comtrade, own graph

Figure 4.3 demonstrates that and enlarging EU has maintained about the same number of trading partners over time in each of the product categories, which means a decline in the number of trading partners relative to the number of EU members. This view is exacerbated when the number of exporting countries is expressed relative to the total volume of imports (right panel).

Apparently, countries have rather specialised in exports to the EU than would new countries have gained market access. One important explanation for this is the EU's system of preferential trade agreements. For instance, it is plausible that the EU as

a major beef and dairy producer has limited interest in beef and dairy imports, while it has substantially less own production of certain tropical fruit. It is therefore necessary to control for different levels of import tariffs applied by the EU when assessing the potential role of institutions in exporting countries.

However, the different number of exporting countries in each product category may also point towards relative differences regarding the ease of implementing agri-food standards that are set by the EU. As long as these standards can, besides tariffs and other trade barriers, be viewed as part of the effect observed in Figure 4.3, then only a relatively small number of countries seems to manage to comply with meat and dairy product standards, while a relatively large number of countries seems to comply with standards for fishery products, fruits and vegetables.

Furthermore, a certain country might in principle be eligible for exports to the EU, but shipments may have been rejected at the border, e.g. due to violations against MRLs. The latter might be especially relevant in the case of fruit and vegetable products.

FIGURE 4.4: Rejections of food imports to the EU.

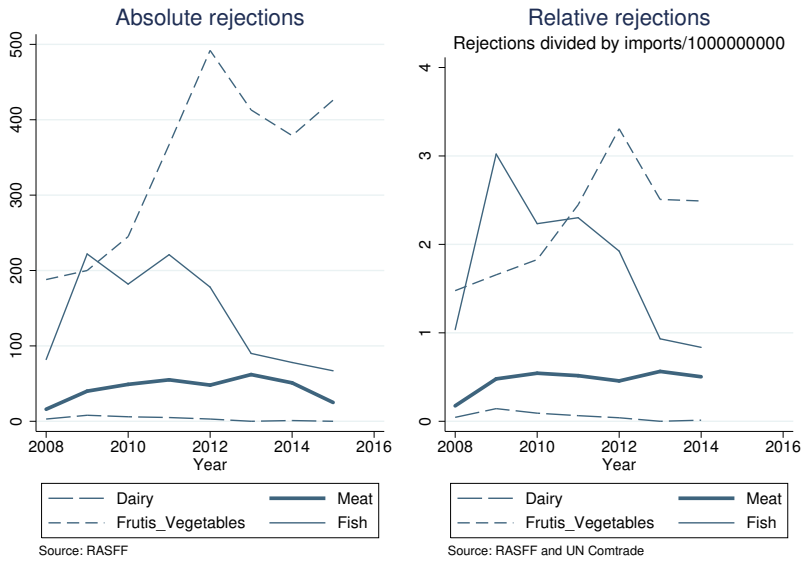


Figure 4.4 provides absolute and relative counts of agri-food shipments that were rejected at EU borders over time. The relative count takes the import value and each year as a reference in order to control for the fact that more frequently traded product categories may also exhibit higher rejection numbers in absolute terms. Figure 4.4 however reveals that absolute and relative rejections in the year 2014 were highest for fruits and vegetables, while dairy and meat products encountered the smallest number of border rejections.

In other words, the findings in Figure 4.4 can be interpreted as the outcome of the two dimensions of agri-food trade standards, namely product quality and process quality: the EU does not restrict so severely the number of countries eligible for the export

of fruits and vegetables a priori, but the type of quality restrictions in place and the corresponding monitoring procedures lead to rather frequent rejections at EU borders.

In turn, the EU seems to strictly limit the number of countries eligible as exporting partners for dairy and meat products, and this may happen through a combination of tariff-related measures and the rather tight monitoring of the entire supply chain within each country that is eligible for exporting these products.

Although institutional quality enhances exports of both quality types — product quality and process quality — it can be expected to be in particular relevant for products which require high levels of process quality, such as meat and dairy products (European Commission, 2017a).

Fishery products constitute in this respect the product category in the middle between fruits and vegetables (as an example of mostly product quality related standards) and dairy and meat products (as examples of mostly process quality related standards): a large number of countries is eligible for exports to the EU, while the number of rejections ranges between fruit and vegetable products on the one hand and dairy and meat products on the other.

4.3.4 Endogeneity of institutional quality

The potential endogeneity of institutions has been emphasised in various studies related to empirical growth and empirical trade analyses (Frankel *et al.*, 1999; Acemoglu *et al.*, 2001; Dollar *et al.*, 2003; Doyle *et al.*, 2011). For example, reverse causality could cause institutions to be endogenous since large trade volumes and historical trade flows could improve the quality of public institutions as well via spillover effects and knowledge transfers.[1]

A valid and relevant instrument is difficult to find. Existing instruments like for example settler mortality (Acemoglu *et al.*, 2001) are often limited to few countries and few – if not only one – years. Using these instruments would reduce the benefits of our rich panel dataset. Other instruments that are primarily applied in growth regressions are the population share of English speaking people (Hall *et al.*, 1999) or geography-based instruments like distance to the equator (Bénassy-Quéré *et al.*, 2007). However, most of them have been criticised as well, for example for weak first-stage results in particular settings.

We address potential endogeneity by using the one-year lag of WGI. At the first place, there is no argument why current trade flows should affect WGI one year ago.[2] This identification strategy has been used in previous studies as well. Dollar *et al.* (2003) for example instrument changes in institutional quality with the one-year lag as explanatory variable in growth regressions.[3]

[1] See Chapter 5 in which we explicitly use these spillover effects and knowledge transfers as arguments for an instrumental variable approach.

[2] We discuss in Chapter 5 under which conditions this assumption is invalid.

[3] However, their empirical approach of using regressions based on growth rates instead of levels has been questioned by (Pritchett, 2003).

4.3.5 Econometric model and estimation procedure

In order to test the hypotheses econometrically, a structural gravity model is estimated using the Poisson pseudo-maximum likelihood (PPML) estimator to account for heteroscedasticity and the high share of zeros (36%) as the two most serious empirical challenges in gravity modelling (Santos Silva *et al.*, 2006; Santos Silva *et al.*, 2011). Although other estimators can be in principal superior under specific circumstances (Martínez-Zarzoso, 2013), the PPML estimator has been widely applied in recent years and therefore constitutes in our eyes the current state-of-the-art.

The PPML estimator does not rely on a linear model, which allows to use trade data in its non-logarithmic form as the dependent variable. Hence, zeros remain in the data set and are not eliminated as the log-transformation requires in linear models. Thus, X_{ijpt} denotes bilateral imports of country j from country i of product p in year t. The final estimation equation is given as:

$$\begin{aligned} X_{ijpt} = \exp(&\beta_0 + \beta_1 \text{WGI}_{it-1} + \beta_2 \text{WGI}_{it-1} * D_{EU} + \beta_3 \text{WGI}_{it-1} * D_{EU} * D_{Product} \\ &+ \beta_4 \ln\text{GDP}_{it} + \beta_5 \ln\text{GDP}_{jt} + \beta_6 \ln\text{Dist}_{ij} + \beta_7 \ln(T_{ijpt} + 1) \\ &+ \beta_8 \Omega_{ij} + \eta_i + \nu_j + \delta_p + \mu_t) + \epsilon_{ijpt} \end{aligned} \tag{4.2}$$

GDPs of the importing as well as of the exporting country are used as proxies for the economic size of both trading partners whereas distance is used as proxy for trade costs. Country-pair and product-specific ad-valorem equivalent tariffs T_{ijpt} are included as well. Furthermore, the matrix Ω_{ij} contains additional dummies as trade costs proxies, namely colonial ties, common language, common border, and religion (compare Table C.4). Importer-, exporter-, product-, and year-fixed effects are included to control for multilateral resistance (Anderson *et al.*, 2003; Head *et al.*, 2014).

Ideally, time-dimensional fixed effects for exporters and importers would also have to be added to the estimation equation in order to control for time-varying multilateral resistance. However, this is computationally not feasible because it would leave the dummies for institutional quality redundant, since they do not vary according to p. Instead of time-dimension fixed effects, we therefore apply the Taylor series approximation of Baier *et al.* (2010) as a robustness check.

In order to test the effect of institutional quality on exporters' ability to implement more demanding standards, it is assumed that the EU constitutes an example of a market with relatively high standards. For this purpose, institutional quality of the previous year of the exporting country is interacted with the dummy D_{EU} (unity for tradeflows towards the EU) and a product dummy for product p.

This specification allows to test whether product-specific effects of institutional quality exist regarding exports to the EU as an example of a market with relatively demanding product standards. The one-year lag of WGI is furthermore used in order to address potential endogeneity due to reverse causality. This specification is also

plausible because the process of implementing a new standard, and obtaining EU approval for it, would typically take several months, if not years.

4.4 Results

Table 4.2 reports the results of four different specifications of the econometric model outlined in Equation 4.2. In columns one and two, importer-, exporter-, product- and year-fixed effects are used to control for multilateral resistance whereas coefficients in columns three and four are obtained by using the Baier and Bergstrand approach including product dummies. Furthermore, both equations have been estimated with and without distinguishing between products through the interaction term of β_3. In column one, the partial effect of WGI, which is the sum of β_1 and β_2 is close to zero which indicates on average no effect of WGI on trade in this specification.

The interpretation of further estimated coefficients is overall in line with the literature of gravity models on trade in general: negative and large effects of tariffs and distance on trade flows are found occurring together with positive effects of income of the importing country as well as from sharing a border.

The dummies for different agri-food products reveal that higher levels of institutional quality (WGI) increase *c.p.* on average exports of dairy products to the EU but are negatively associated with exports of fruits. Exports of meat, vegetables, and fish are not affected by WGI. In contrast to column one, the Baier and Bergstrand approach in column three reveals a partial effect of WGI on exports in general, but to the EU in particular. In addition, the coefficient of tariffs declines from -2.261 to -1.080. The coefficient of the exporting country's GDP becomes statistically significant, and so do the coefficients for former colony and religion, respectively. These findings are complemented by a positive and statistically significant effect for dairy, meat, and vegetable products. The coefficients for exports of fruits (negative), and fish (not statistically significant) remain as in the model in column two.

Thus, results overall suggest a trade-increasing effect of WGI in particular to the EU as a market (that we have used as an example of a market with relatively high trade-related standards) for the agri-food products in question. This implies that hypothesis one cannot be rejected based on the results in Table 4.2.

However, the overall positive effect of institutions on agri-food exports to the EU cannot be generalised across product categories. Instead, WGI turns out to be especially relevant for exports of dairy and meat products. These products of animal origin provided examples of mostly process quality related standards. The effect of governance indicators on the export of fish and fish products to the EU is insignificant throughout, and the positive and significant effect of WGI on vegetable exports appears in only one of two model specifications. For fruits, the observed effect is even negative, implying that countries with relatively well-functioning public institutions according to the WGI are less likely to export fruits to the EU.

Table 4.2: Regression results - dependent variable: Imports

VARIABLES	(1) PPML	(2) PPML	(3) PPML-BB	(4) PPML-BB
WGI_1	-0.0291*	-0.0202	-0.00516	-9.34e-05
	(0.0173)	(0.0144)	(0.00740)	(0.00756)
WGI_EU_1	0.0230	0.00955	0.132***	0.0112
	(0.0288)	(0.0157)	(0.00984)	(0.0132)
WGI_EU_Dairy		0.311***		0.425***
		(0.0512)		(0.0151)
WGI_EU_Fruits		-0.162***		-0.0725***
		(0.0440)		(0.0133)
WGI_EU_Meat		0.103*		0.179***
		(0.0573)		(0.0160)
WGI_EU_Vegetables		0.0483		0.148***
		(0.0480)		(0.0170)
WGI_EU_Fish		-0.0748		0.00868
		(0.0498)		(0.0167)
lnAdvalorem_Tariff	-1.870***	-2.261***	-0.672***	-1.080***
	(0.632)	(0.626)	(0.198)	(0.202)
lnGDP_Exporter	0.0500	0.0466	0.384***	0.394***
	(0.0408)	(0.0443)	(0.00539)	(0.00539)
lnGDP_Importer	0.887***	0.876***	0.692***	0.697***
	(0.0449)	(0.0460)	(0.00525)	(0.00528)
lnDist	-0.800***	-0.803***	-0.708***	-0.700***
	(0.0469)	(0.0466)	(0.0127)	(0.0124)
Contig	0.699***	0.690***	0.671***	0.671***
	(0.137)	(0.135)	(0.0435)	(0.0431)
Comlang_off	0.102	0.105	-0.197***	0.207***
	(0.123)	(0.121)	(0.0357)	(0.0351)
Colony	0.0954	0.0883	0.131***	0.158***
	(0.122)	(0.123)	(0.0346)	(0.0338)
Constant	-11.24***	-11.00***	-13.31***	-13.74***
	(1.479)	(1.659)	(0.222)	(0.223)
Observations	512,257	511,055	512,257	511,055
R-squared	0.369	0.410	0.261	0.297

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This finding is interesting because the gravity equation has already controlled for the fact that some former colonies of EU member states have traditionally strong trade relations to the EU. Some of these countries at the same time happen to be both exporters of tropical fruit and examples of particular poor public institutions. Beyond that, the results suggest that rather low scores on the WGI may still constitute a source of comparative advantage for the export of fruits, but not vegetables.

From the perspective of the EU food inspection agency, this result might potentially be worrisome, given that one explanation could be that for major food products certain stages of the approval process could be subject to corruption. Alternatively, the result may point that socioeconomic circumstances such as market power along fruit chains, oppression of seasonal farm workers or other issues that may be more widely spread under poorly functioning public institutions while at the same time leading to more competitive exports of fruits.

Thus, regarding our second hypothesis, it is difficult to argue whether higher SAA indeed favours exporting those agri-food products for which standard implementation is relatively more demanding than for other agri-food products. Instead, our findings show that especially the process quality -related standards in case of dairy and meat products clearly benefit from better public institutions, and this may indeed contribute to higher SAA. The negative coefficient on fruit exports however shows that in reality, this effect may for specific agri-food products be dominated by export competitiveness that is rather due to especially low quality public institutions.

4.5 Discussion and Conclusion

We find an overall positive effect of institutional quality on countries' exports to the EU as a market with relatively strict standards. This implies that our hypothesis according to which institutional quality facilitates firm's ability to solve export related problems cannot be rejected. Furthermore, once interaction terms with dummies that mark product-specific exports are introduced, our results show that institutional quality is only trade increasing for dairy products and meat, as examples of product categories with rather process quality related standards.

In contrast, the rather product quality related standards in the area of fish, fruits and vegetable exports do not confirm this finding. Dairy and meat products tend to show more processing stages and more intensive monitoring through the EU and more difficult approval procedures for producing firms and processing factories in the exporting country than for the case of fruits and vegetables on average.

We interpret these results as empirical support for our second hypothesis, according to which standard implementation problems tend to be more difficult to solve for firms in product categories that exhibit more complex steps of the corresponding supply chains. However, our evidence regarding the second hypothesis is more mixed than concerning the first one.

Regarding predictions of the New New Trade Theory, these findings suggest that two firms with otherwise identical productivity that are producing identical product varieties but having to operate in two different institutional environments, will face different probabilities of succeeding in the implementation of EU standards.

If one assumes that the EU is an example of presumably rather higher than average product standards, our analysis furthermore suggests that even within the agri-food sector the effect of institutions on export performance varies according to the type of product. This is according to our analysis due to the fact that market access is (in terms of required problem-solving capacity) comparatively difficult to achieve for exporters in agri-food products of higher processing order. We explain this finding as the result of trade-related standards for these products being relatively more knowledge-intensive and therefore relatively more affected by a lack of institutional quality in the exporting country than is the case for less knowledge intensive products.

With respect to the discussion of standards as trade catalyst versus trade barriers, our results reject any across the board conclusion and rather highlight that these effects will be highly product specific in therefore inseparable from the specific agri-food supply chain in question. Regarding the fixed cost character of standards, our results show that it is necessary to interpret these fixed cost not only in terms of required investment for laboratories and processing plants, but also in terms of private sector human capital and public sector institutional quality.

Especially the latter two have according to our results to go hand-in-hand in order to create not only necessary but sufficient conditions for agri-food exporting firms to succeed in high quality and potentially high-value export markets.

Future research needs to extend the list of included products. Moreover, other destination markets with relatively strict standards such the USA or Japan would be worthwhile to analyse. More robustness checks regarding endogeneity of WGI are also required. Although we do control for income, it would be interesting to know to what extent these findings differ by the income level of the exporting country.

Due to the diverse effects of WGI on exports we conclude that aid-for-trade programmes should focus on building up production capacities of agricultural products which are relatively complex to produce. Although we do not distinguish between the extensive and intensive margin of trade, our findings suggest that improving the quality of public institutions could increase the relatively low number of exporting countries of complex products. Hence, this discussion relates to the conclusion of the previous Chapter 3 in the sense that a sufficient number of exporting countries is desirable to avoid potential market failures due to too concentrated markets. In other words, higher quality of public institutions would not only increase exports of these products but also reduce the risk of too concentrated agricultural export market by making the integration of countries into global value chains more likely.

To what extent (private) standards are a useful policy tool to increase exports of relatively complex agricultural products, is subject of the analysis of the subsequent

Chapter 5. It complements findings of this chapter because it investigates the effect of an important private food standard on exports of processed food standards.

5 The role of private standards for manufactured food exports from developing countries

Abstract:

The relevance of non-tariff barriers for global trade flows has increased in recent decades. However, the effect of food standards – as a particular important non-tariff measure – on agricultural trade flows remains unclear. We contribute to the debate with a unique dataset that contains the number of food processing firms of 88 countries from 2008 to 2013 that are certified with the International Featured Standard (IFS). We estimate a gravity-model using the one-year lag of IFS as well as IFS certification in neighbouring countries as an instrument to address potential endogeneity. We find that IFS increases *c.p.* bilateral exports on average of seven agricultural product categories in both specifications. The effect is more pronounced for high-income countries compared to low income countries. Hence, whereas IFS increases exports on average, the trade-enhancing effect remains only for high-income countries. We argue that food standards are not necessarily a suitable development tool to integrate low-income countries into high-value chains per se because high-income countries benefit more in terms of trade flows than low-income countries. Moreover, we contribute to the debate about food standards and trade by specifying the nature of IFS and – based upon our theoretical model of Chapter 2 – under which circumstances food standards enhance or reduce agricultural trade.

This chapter is joint work with Axel Mangelsdorf. An earlier version is under revision at *World Development*.

Malte Ehrich's contributions are: Development of research question and research design including instrumental variables, empirical and theoretical framework, empirical estimation, and writing of the manuscript. Axel Mangelsdorf provided the dataset, excluding instrumental variables and tariffs.

Acknowledgements: We are thankful for comments of Marie-Luise Rau and other participants of the ETSG conference in Helsinki 2016, and for valuable comments of Bernhard Brümmer, Miet Maertens, Insa Flachsbarth, and Katharina Krumbiegel.

5.1 Introduction

Chapter 1 already emphasises the increasing relevance of public but also of private food standards for international trade. While public standards, set by public authorities, are usually mandatory and legally enforceable, private standards are set by the private sector which often have a wider scope than only food safety (Schuster *et al.*, 2015) and are on average more restrictive (Fulponi, 2006). As a result, the effect of private standards on trade is likely to differ from the effect of public standards.

The debate of "standards-as-catalyst and standards-as-barriers" has been described in Section 1.2. The different types of standards, like public vs. private, are likely to constitute one source for the heterogeneous findings in empirical literature. Therefore, this chapter contributes to the debate by analysing the effects of IFS on trade as one particular and important private standard for processed food products.

For example, compliance with private standards can be more costly due to the higher degree of restrictiveness on average (Fulponi, 2006) which would support the trade-barriers view. In contrast to MRLs (see Chapter 3), private standards affect trade via additional channels: first, private standards might be more capable to signal high product quality and hence, might be more likely to meet preferences of modern consumers. If small-scale farmers are given the opportunity to enter global value chains by certifying their products with GlobalGAP, Fairtrade, or other certificates, they are more capable to convince western consumers about the quality of their agricultural products. Second, private standards might reduce transaction costs of producers and within value chains overall (David *et al.*, 1990). This is in particular relevant for IFS. It is an important B2B-standard that was established by German and French retailers to harmonise standards within value chains. This objective of western retailers explicitly addresses transaction costs. Thus, compliance with IFS is likely to reduce transaction costs and hence, facilitates integration within value chains which finally *c.p.* increases trade flows. Therefore, we expect IFS to increase trade flows in contrast to other standards like MRLs.

The potential ability of private standards to integrate small-scale farmers from developing countries into global value chains makes them a complementary tool for development strategies (Swinnen, 2007; Swinnen *et al.*, 2015; Colen *et al.*, 2012). Since standards potentially reduce market failures due to lower information asymmetries, food standards might be more relevant for developing countries in particular (Jaffee *et al.*, 2004). If private food standards are found to increase exports of developing countries, this would have important policy implications. In addition to the poverty reducing effect due to larger trade volumes, food standards would facilitate equal access to global agricultural export markets. The latter is important from a global perspective since it improves the functionality and benefit of global markets for all participants.

Moreover, there are potential additional benefits at the firm level. Since trade is not only welfare-enhancing via lower consumer prices, export sectors are on average also the most competitive sectors in a country. Thus, exporting firms earn on average

higher profits, employ a larger number of workers, and pay higher wages than non-exporting firms worldwide (Mayer *et al.*, 2007). Colen *et al.* (2012) provide empirical support that this pattern occurs in developing countries as well. In the context of GlobalGAP certification in Senegal, the authors show that exporting firms are important drivers for job creation and productivity spillovers which underlines the potential of private food standards as a development policy tool.

Because of their private nature, data on private standards are more difficult to obtain than for public standards which are often publicly available. As a result, private standards are less frequently analysed. Using firm-level data of the Peruvian asparagus sector Schuster *et al.* (2015) cannot confirm that BRC, IFS, and other private standards act as catalyst to trade. Although private standards are on average more stringent than public standards (Fulponi, 2006), these have nevertheless the potential to increase agricultural trade. Masood *et al.* (2014) find that GlobalGAP certification increases banana imports of the EU. The trade enhancing effect of GlobalGAP certification is also found by Colen *et al.*, 2012 for mango and bean producers in Senegal which have larger export market shares and larger export volumes than non-certified firms. The differential effect of voluntary private standards compared to public standards on trade is also emphasised by Shepherd *et al.* (2013) who find that EU harmonised standards, that are equivalent to ISO norms, can even enhance trade. Eventually, Mangelsdorf *et al.* (2012) estimate the effect of Chinese public and private standards and also find a trade-enhancing impact. The latter was most pronounced for internationally harmonised standards.

Overall, few studies exist that analyse the effect of private food standards on agricultural trade. Moreover, existing analyses use either cross-sectional data (Latouche *et al.*, 2015) or are based on data which are limited to specific cases. This questions external validity as emphasised by Beghin *et al.* (2015). Major challenges are the quantification of private standards and data availability. Most studies do not allow to draw general conclusions because they are based on very few products and countries. Furthermore, endogeneity of the standard variable arises as a result of reverse causality. A correct identification of the causal impact requires to distinguish whether it is certification that enhances trade or whether trade increases the likelihood of certification. Finally, a correct specification of the empirical framework requires to account for recent developments in the field of gravity modeling which became the workhorse model in empirical trade analysis (Head *et al.*, 2014).

We address these shortcomings, first, with a unique dataset which was obtained via the IFS auditing database. In contrast to previous studies on private standards and trade, the dataset is rich in all dimensions: It contains more than 50,000 audits from about 12,000 companies in 88 countries for seven agricultural product categories including a time-span of six years from 2008 to 2013. Second, we apply a novel instrumental variable approach which we consider to be superior compared to the standard method of taking a one-year lag which is not appropriate if the errors are autocorrelated. Third, we estimate a gravity model via PPML which accounts for high share of zeros and heteroskedasticity (Santos Silva *et al.*, 2006; Santos Silva *et al.*, 2011). Furthermore, we apply the Baier-Bergstrand method to address multilateral resistance (Baier *et al.*, 2010; Anderson *et al.*, 2003). This approach allows us

to contribute to the debate whether standards act as barriers or catalyst to trade.

We find that IFS certification as a private standard increases bilateral trade flows in general which illustrates the trade increasing potential of IFS. However, the effect remains robust only for high-income countries while effects are unclear for low-income countries once we distinguish by income groups. This finding has important policy implications. Although IFS certification increases trade on average, only high-income countries benefit in contrast to low-income countries. This finding reduces the potential of food standards as a development tool to integrate developing countries into the world trade system.

The remainder of the paper is structured as follows: Section 5.2 provides additional informational background on IFS certification. The subsequent Section 5.3 explains the PPML-estimation and the instrumental variable approach in particular including the control-function approach. Section 5.4 shows the results, which are discussed within the research context in Section 5.5.

5.2 IFS background and trends

The increasing complexity of agricultural value chains due to fragmentation and specialisation increases the necessity for sufficient transparency within value chains. Retailers need to guarantee quality and food safety of the products that they sell, but which they do not produce themselves. Moreover, to ensure the enforcement of legal contracts it is crucial to have transparent responsibilities at every stage within a value chain. Therefore, the association of the German retail sector HDE[1] found together with the French counterpart FCD[2] the initially named International Food Standard in 2003. The IFS is applicable at every stage of a value chain apart from agricultural raw products. This private standard – today the International Featured Standard – avoids that each retailer is required to test whether their suppliers meet the imposed standards or not. Instead, retailers agreed on the same standards. These standards are continuously modified in collaboration with the retail sector. Hereby, most regulations go beyond usual food safety standards (International featured standard, 2016). The overall objectives are twofold: first, IFS ensures comparability, transparency, and quality for the consumer within a complete value chain. Second, it aims to reduce costs for the retail sector and their suppliers by harmonizing standards. Apart from the UK, where the BRC is the most relevant food standard certification body, all major retailers within Europe are member of the IFS[3] which also certifies in other fields like logistics for example.

IFS does not certify products and food manufactures directly but rather via third party certification bodies which takes place on average once a year. All retailers that accept the IFS have access to these audit reports of their suppliers via an online database. In addition, all certified producers have access as well. But apart, access

[1] Hauptverband des Deutschen Einzelhandels

[2] Fédération des Entreprises du Commerce et de la Distribution

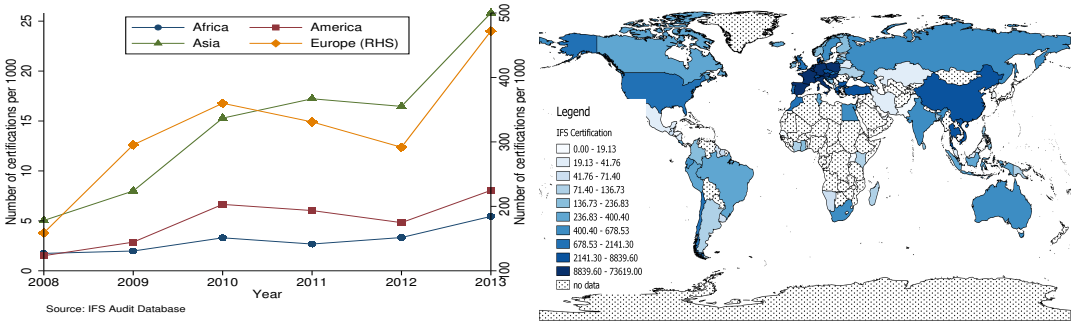
[3] Metro Group, Edeka, Rewe Group, Aldi, Lidl, Kaufland, Kaiser's Tengelmann, Auchan, Carrefour Group, EMC – Groupe Casino, Leclerc, Monoprix, Picard, Surgelés, Provera (Cora and Supermachés Match), Système U, COOP, CONAD und Unes.

to the database and information concerning audit reports and other confidential data is not possible.

Moreover, it is not only the availability of IFS data that makes our analysis distinct from previous studies. As a post-farm gate standard, which needs to be distinguished from pre-farm gate standards like GlobalGAP that certify agricultural raw products, IFS certifies processed food. These manufactured food products yield a higher value-added than non-processed food products. Hence, certification with IFS is expected to generate even higher profit than other standards (Colen *et al.*, 2012).

We consider our data set as unique since it contains the number of certificates per country and product from 2008 to 2013. The amount of certification is an indicator for the relevance of IFS within a country. The about 12,000 food manufacturing companies are located in 88 countries including 53 developing countries.[4] The total number of certification increased from about 4,000 in 2008 to almost 12,000 certificates in 2013. Europe is the major hub of IFS certification, see Figure 5.1. Numbers increased especially in Asia and Europe from 2008 to 2013 by almost 500% and about 100% respectively. The unequal distribution of IFS with Europe being the most relevant region is displayed in the world map in Figure 5.1 as well.

FIGURE 5.1: Regional aspects of IFS (1)

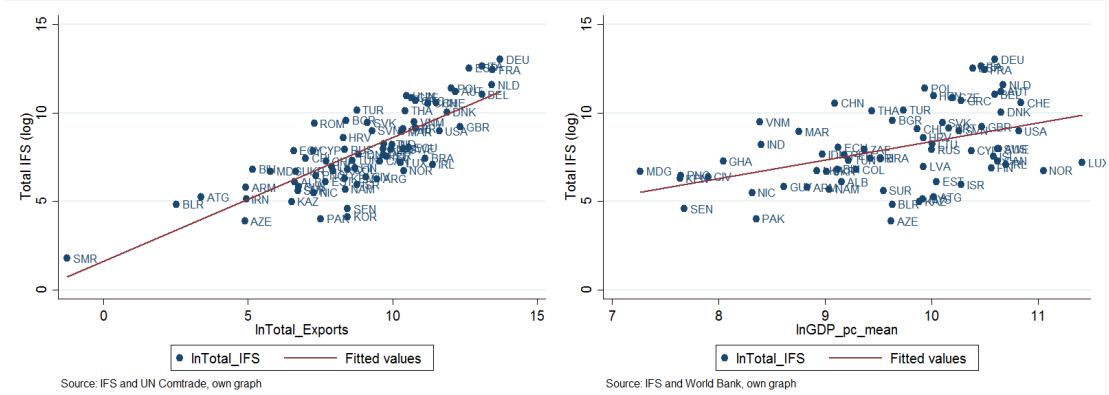


(a) Numbers of certificates per continent (b) Number of certificates worldwide

Moreover, Figure 5.2 underlines some regional patterns regarding the correlation between exports and income. Countries with more IFS certification tend to have higher exports with central European countries leading both IFS and exports. Similarly, richer countries have more certified producers on average. Both patterns naturally reflect the dominance of central European countries.

[4] All products and countries including HS classification are listed in Tables D.1 and D.2 respectively.

FIGURE 5.2: Regional aspects of IFS (2)



(a) IFS and exports

(b) IFS and per capita income

5.3 Model specification

5.3.1 Data

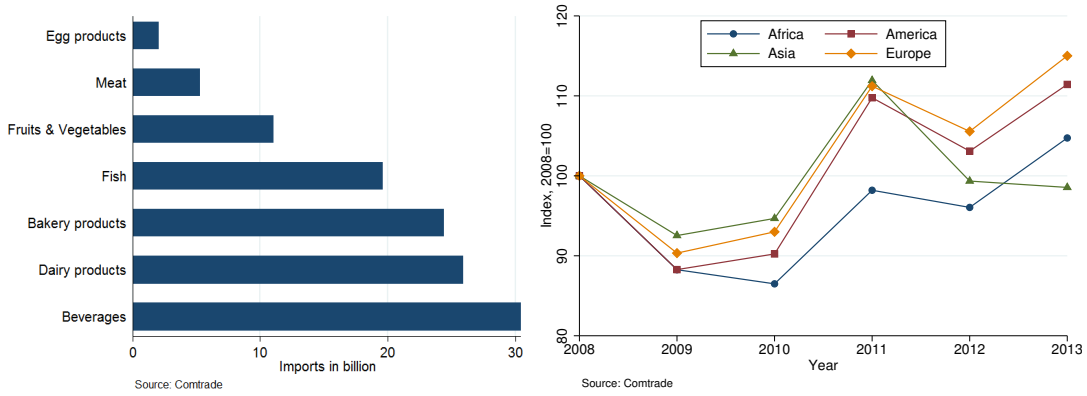
Since we are predominantly interested in the effect of IFS as an important private standard on bilateral trade flows, IFS is our main variable of interest. As highlighted by Head *et al.* (2014), the gravity model became the "workhorse model" in empirical trade analysis. The required variables are explained briefly:

Bilateral trade in current US Dollar from UN Comtrade is the dependent variable for seven different product categories: egg products, meat, fruits and vegetables, bakery products, dairy products, and beverages. We use seven importing countries in which IFS is widely applied: Austria, Belgium, France, Germany, Italy, Netherlands, and Switzerland, see Table D.2. In addition to total export values per product, Figure 5.3 also displays the export performance per continent. Hereby, we use an index equal to 100 for the year 2008. Exports declined for all four continents until 2009 due to the economic and financial crisis, peaked in 2011 and mostly increased again for 2013. Asia is the only continent that performed worse in 2013 compared to 2011. Exports overall increased by 11% and Europe is the best performing continent with increasing exports of 15%.

The remaining variables are of standard gravity nature: we include the logarithm of GDP in current US-Dollar from the World Bank as proxies for the economic mass of both trading partners. Proxies for trade costs like distance, language, and colony are obtained from CEPII whereas ad-valorem tariffs come from the ITC. Descriptive statistics are provided in Table 5.1. In total, we use 146,091 observations from which 58% of the deflated export observations are equal to zero.

Since IFS was particularly designed by modern retailers and therefore, for high value chains, we expect IFS to have high values of θ - the "ease" to address preferences of modern consumers. Thus, compliance with IFS allows producers to sell their products at high-value markets. We expect IFS to allow producers to make their products

FIGURE 5.3: Exports per product and continent



(a) Exports in 2013 / product categories

(b) Export performance (2008=100)

TABLE 5.1: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
GDP_Exporter (WDI, current US-D, logs)	146,091	25.207	1.967	20.563	30.305
GDP_Importer (WDI, current US-D, logs)	148,131	27.520	.843	26.491	28.758
Distance (CEPII, logs)	148,131	8.044	1.085	4.088	9.740
Exports (UN Comtrade)	148,131	4,324.555	29,524.54	0	1,343,673
Exports (UN Comtrade, deflated)	146,091	37.519	263.816	0	12509.41
RTA (CEPII)	148131	.2424611	.4285731	0	1
Language (CEPII)	148131	.0847291	.278479	0	1
Colony (CEPII)	148131	.0332	.180	0	1
Tariff (ITC, logs)	148131	.052	.136	0	1.643
IFS certification (IFS Audit database)	148,131	13.829	52.373	0	694

distinct from others. This is a crucial difference to public mandatory standards like MRLs since these need to be met in any case. Based on this argumentation, we expect IFS to increase trade. However, compliance costs might still be too high such that developing countries might not be able to benefit from IFS as an opportunity to gain access to high-value markets. Hence, we expect differential effects of IFS on trade depending on the income level of the country of origin.

5.3.2 The benchmark specification

The estimation strategy of gravity models in international trade needs to address several empirical challenges. The model needs to account for multilateral resistance (Anderson *et al.*, 2003), high share of zeros (Helpman *et al.*, 2008), and heteroskedasticity (Santos Silva *et al.*, 2006; Santos Silva *et al.*, 2011) in particular. Country-year fixed effects are frequently used to account for multilateral resistance. However, this approach becomes computationally difficult the larger the data set becomes in terms of countries and years included. Alternatively, Baier *et al.* (2010) propose a different method which adjusts all trade cost proxies in such a way that multilateral resistance does not differ across countries. We chose country-, year-, and product fixed effects as our baseline specification as well as the Baier-Bergstrand method. Mainly because country-year fixed effects are computationally difficult but also because multilateral resistance is less likely to change over time during the relatively short time period of five years, we expect country- and year fixed effects to capture multilateral resistance well. Moreover, we estimate a multiplicative gravity model with PPML (Santos Silva *et al.*, 2006) which does not require to take logs of the dependent variable and therefore, does not drop zeros. In addition, it is robust to heteroskedasticity which is usually present in trade data. The final model of the benchmark specification is defined as follows:

$$\begin{aligned} X_{ijpt} =& \exp(\beta_0 + \beta_1 \ln\text{IFS}_{ipt-1} + \beta_2 \ln\text{GDP}_{it} + \beta_3 \ln\text{GDP}_{jt} + \beta_4 \ln\text{Dist}_{ij} \\ &+ \beta_5 \ln\text{Tariff}_{ijpt} + \beta_6 \text{Language}_{ij} + \beta_7 \text{Colony}_{ij} + \beta_8 \text{Contiguity}_{ij} \\ &+ \beta_9 \text{RTA}_{ijt} + \mu_i + \nu_j + \lambda_t + \upsilon_p)\eta_{ijpt} \end{aligned} \tag{5.1}$$

X_{ijpt} denotes deflated exports from country i to country j of product p in year t. IFS represents the number of certifications in the exporting country and is the main variable of interest. However, IFS is likely to be endogenous because of reverse causality. Certification might not only increase trade flows due to the beforehand explained reasons. Vice versa, products might be more likely to be certified if trade flows are high. Therefore, in the benchmark specification IFS is introduced as a one-year lag to address partially endogeneity due to reverse causality.

5.3.3 An instrumental variable approach

Because of the above mentioned reverse causality, we expect IFS to be endogenous. The lag of IFS as an instrument for IFS does not solve the endogoneity problem if the errors η_{ijpt} are autocorrelated.

$$E\left(\text{IFS}_{ipt}\eta_{ijpt}\right) \neq 0$$

If the one-year lag of IFS was exogenous, it should not be correlated with the error term:

$$E\left(\text{IFS}_{ipt-1}\eta_{ijpt}\right) = 0$$

This argument is based on the assumption that IFS itself is correlated over time but the errors are not. However, in the presence of autocorrelation the one-year lag is not a valid IV:

$$\eta_{ijpt} = \rho_1\eta_{ijpt-1} + \iota_{ijpt}$$

If the coefficient $\hat{\rho}_1$ is significantly different from zero, the error terms are autocorrelated and the exogeneity assumption of IFS_{ijpt-1} does not hold. If IFS is correlated with the current error term, it is also correlated with its lag if the errors are autocorrelated. The Wooldridge test for autocorrelation rejects the null hypothesis of no autocorrelation for all usual significance levels. Therefore, an additional identification strategy is required.

IFS certification in neighbouring countries as an instrument

Applying an instrumental variable approach allows to address the endogeneity of IFS *if* a valid instrument is available. Relevance and excludability are the two key requirements of a valid instrument. An instrument is relevant if it explains sufficient variation of the endogenous variable. This relevance-condition can be tested empirically for example via partial R-squared of the first-stage estimation. In contrast, excludability (or strict exogeneity) as the second condition is not testable and requires arguments based on economic theory. An instrument is excludable if it affects the outcome variable *only* through the endogenous variable. This requirement makes it difficult and sometimes impossible to find a valid instrument. As we will argue in the following, the total number of IFS certified producers in all neighbouring countries of a particular exporting country i meets both requirements.[5]

This instrument has been used previously in the context of genetically modified organisms (GMO) regulation and trade (Vigani *et al.*, 2012). The authors use the weighted average of GMO indices of the five closest neighbours to avoid biased estimates due to endogeneity. The estimated negative effect on trade due to GMO regulation even increases in magnitude compared to the non-IV specification. In a similar approach, Djankov *et al.* (2010) use export delay of neighbouring countries as an instrument of

[5] We used modern grocery distribution as an instrument in an earlier version of the paper. However, the prevalence of modern grocery stores is correlated with income which is – by definition of gravity – a determinant of trade. As a result, we make use of a different instrument which is explained below.

domestic trade time.

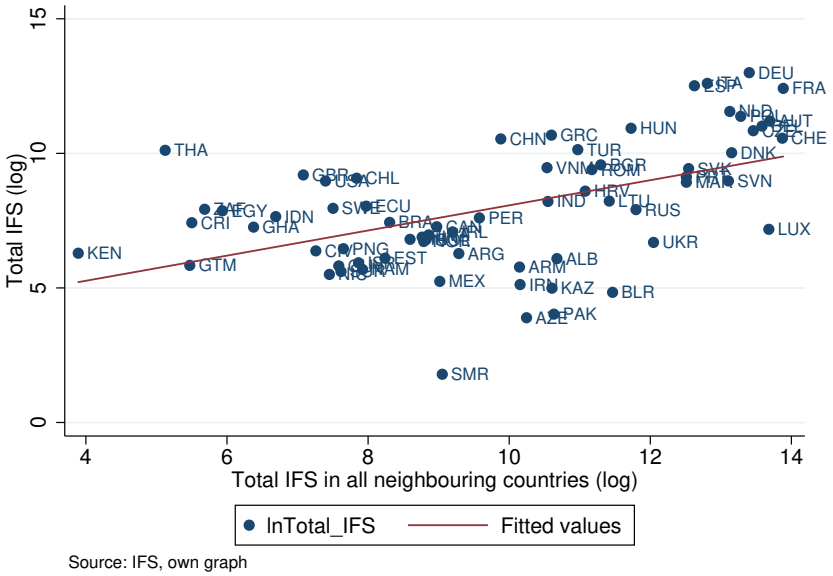
Using specific characteristics of neighbours as instruments is well established in particular at the micro-level. For example Chege *et al.* (2015, p.398) use the "number of supermarket farmers among the five nearest neighbours" as an instrument for supermarket participation of a particular farmer. Based on Maertens *et al.* (2013), the authors argue that social interactions within a neighbourhood of farmers determine adoption behaviour of modern agricultural technologies. These effects and mechanisms are also present and verifiable at the macro-level. There is empirical evidence for the relevance of neighbours for exports of a particular country. Kamal *et al.* (2016) show that firms of an exporting country i are more likely to establish trade relationships with firms in an importing country j if firms of a neighbouring country of i have already well established relationships with firms in j. The authors discuss various channels like knowledge transfer in terms of destination-specific cultural and business norms as well as legal requirements. A second channel refers to cost-sharing like reducing search- and matching costs between sellers and buyers. Both channels are particularly relevant in the presence of imperfect information. These spillover-effects of neighbouring exporters are especially important for the exporting decision (extensive margin), but less relevant for export volume per firm (intensive margin) (Koenig *et al.*, 2010). The authors refer this difference to the varying relevance of fixed- and variable trade-costs. Thus, spillover-effects due to exporting neighbours reduce barriers to trade which are a result of fixed costs. This makes these finding interesting for the debate of standards and trade. Moreover, the effect is stronger at disaggregated levels like product- and destination-dimensions than at the aggregate level. Summing up, there is empirical support in related literature that export decisions of firms within a country are influenced by firms in neighbouring countries.

Networks between these firms facilitate knowledge transfer which we could directly link to compliance with standards like IFS. As shown in Chapter 4, not only monetary costs determine compliance but also non-monetary costs which emphasises the important role of knowledge transfers. Thus, we argue that the extent of compliance with a particular standard in a country affects the likelihood of compliance in a neighbouring country.

Figure 5.4 supports this argument. It shows the positive correlation between the endogenous variables IFS certification and the instrument which is defined as the sum of IFS certification in all neighbouring countries. The correlation is most pronounced in Western European countries like Germany, France (the initiators of IFS) but lower in Easter European countries like Ukraine, Belarus, Albania, and Armenia. However, note that some countries do have neighbours for which we do not have any IFS observation. Nevertheless, the correlation between both measures is equal to 0.51. To conclude, we are confident that the instrument is relevant and therefore, meets the first requirement.

The exclusion restriction requires the instrument to affect exports only via the endogenous variable. In contrast to the benchmark specification, reverse causality is not a problem in this IV specification. There is no plausible argument why exports of a country should influence compliance with IFS in neighbouring countries. Furthermore, there is also no channel through which IFS in neighbouring countries could

FIGURE 5.4: Scatter plot of number of IFS certification in exporting country i and the sum of IFS certified producers in all neighbouring countries.



Source: IFS, own graph

affect exports apart from IFS of the country itself. The extent of IFS certification in a country does neither affect for example GDP of its neighbour nor tariffs or other determinants of exports of their neighbour. Instead, domestic borders are exogenous such that countries cannot self-select themselves towards specific neighbours.

The control-function approach

We modify our estimation strategy of Section 5.3.2 by including an instrumental variable via the "control function approach" as proposed by Wooldridge (2010) and Martínez-Zarzoso (2015).

IFS is the endogenous variable whereas the vector $\mathbf{Z}$ denotes all exogenous variables from which $\mathbf{Z}_1$ is a sub-vector. These are the standard gravity variables like GDP, distance and other trade cost proxies.

$$X_{ijpt} = \mathbf{Z}_1\delta_1 + \alpha_1 \text{IFS}_{ipt} + u_{ijpt} \tag{5.2}$$

The exogeneity assumption can be expressed as follows:

$$E(\mathbf{Z}_1' u) = 0 \tag{5.3}$$

Consequently, the reduced form for IFS is:

$$\text{IFS}_{ijpt} = \mathbf{Z}\pi_2 + \epsilon_{ijpt} \tag{5.4}$$

where $\mathbf{Z}$ includes $\mathbf{Z}_1$ as well as IFS of neigbours as an instrument. The residuals $\hat{\epsilon}_{ijpt}$ of this first stage regression are needed for the linear projection of the residuals of Equation 5.2 u_{ijpt} on these residuals of the first stage ϵ_{ijpt}.

$$u_{ijpt} = \rho_2 \epsilon_{ijpt} + \phi_{ijpt} \tag{5.5}$$

Ultimately, we plug Equation 5.5 into Equation 5.2 to obtain the final control function:

$$X_{ijpt} = \mathbf{Z}_1 \delta_1 + \alpha_1 \text{IFS}_{ijpt} + \rho_2 \epsilon_{ijpt} + \phi_{ijpt} \tag{5.6}$$

If the estimate $\hat{\rho}_2$ is significantly different from zero, we can conclude that IFS is actually endogenous. If IFS was exogenous, there would not be any variation within the reduced form residuals that explain variation of exports in the control function because variation of IFS is completely explained by the vector of exogenous variables $\mathbf{Z}$. As robustness, we also estimate the IV approach with General Methods of Moments.

5.4 Results

We estimate the effect of IFS on trade by using three different methods: first, ppml without IV and two models with IV which are estimated via the control function approach and GMM. Furthermore, we estimate these models at the aggregate level (Table 5.2), by income group of the exporting country (Table 5.3), and finally by product at the sectoral level (Table D.3 in the appendix). This section presents all results.

Column one of Table 5.2 shows coefficients of the benchmark specification which is estimated via ppml and the one-year lag of IFS certification. The coefficient is equal to 0.568 is interpreted as elasticity and it is statistical significant. Thus, a one per cent increase of IFS certification increases country i's exports *c.p.* by 0.568 per cent on average. Other elasticities match general findings obtained via ppml within the gravity framework. Coefficients of GDP-variables are statistically significant and below unity. In contrast, distance and tariffs reduce trade whereas other trade costs proxies language, colonial relationship are contiguity are trade enhancing. The coefficient of regional trade agreement is equal to 1.4 which implies an economic magnitude of 300%.[6] An increase in exports of 300% if a regional trade agreement is signed, is not plausible.

Column two contains estimates of the IV regression via the control function approach. Most are remarkably similar compared with the non-IV method. The IFS-coefficient is about same magnitude equaling 0.518. The few differences are that tariffs and language become statistically insignificant. Moreover, the coefficient of RTA, which was unreasonably high in the benchmark specification in column one, is now equal to 0.815. This implies that is expected to increase by 122% which is still high nevertheless. Column three and four belong to the control function approach. Column three shows estimates of the first stage which indicates the indeed high relevance of IFS

[6] Dummies are interpreted in ppml regressions as $\left(e^{\beta_D} - 1\right) \times 100$.

neighbouring countries for IFS certification itself. The statistical significant constant ρ_2 in column four means that IFS certification is indeed endogenous, see Equation 5.5. Finally, GMM results are shown in column five. Again, most coefficients remain similar. The effect of IFS on trade is slightly higher now (0.635). Summing up, IFS is statistically and economically significant in all three specifications. IFS-certification increases exports *c.p.* on average by about 0.5 and 0.6 per cent if certification increases by one per cent.

However, results differ by income of the exporting country. Table 5.3 shows coefficients of the benchmark specification (columns one, five, and eight) as well as of the discussed IV specifications. We distinguish between high-income, upper-middle income, and low-income groups (see Table D.2). The effect of IFS on trade is positive and statistically significant in the benchmark specification for all income groups. The effect is largest for high-income countries and lowest for upper-middle income countries. However, the effect remains similar in IV models for the high-income group but switches in sign for the upper-middle income group. In fact, the coefficient equals -1.649 which is extremely large. A GMM estimate is not available here due to convergence problems. Moreover, IFS is found to have no effect on exports of low-income countries according to both IV models. Coefficients of distance are unusually large. First-stage results are large in particular for high-income and upper-middle income group which underlines the high relevance of the IV.

Because our dataset also allows to distinguish by sector, we estimate all estimations also at the sectoral level (Table D.3). Due to convergence problems, IV estimates based on GMM are not available for eggs and dairy-products. IFS increases *c.p.* exports of all seven product categories apart from fish where results are mixed. Exports of bakery products increase in particular due to IFS certification. Estimates of beverages, egg, meat, and FV products remain mostly between 0.4 and 0.6 which is of similar magnitude as at the aggregate level. Only the GMM estimator of FV is higher with a value of 0.8. According to the control function approach, exports of dairy products increase drastically as IFS certification increases. The negative coefficient on fish exports in the control function approach remains to be discussed (see Section 5.5 below). The GMM model shows the expected positive sign of reasonable magnitude.

5.5 Discussion and conclusion

After the plain description of the results, this section discusses and interprets the obtained findings. Several aspects ask for deeper discussion.

Interpreation of negative standards coefficients

First, the IFS coefficients for exports of fish is negative in the control function approach. Although the coefficient is positive in the GMM-model which makes a final interpretation difficult, it remains worth discussing what a negative coefficient would mean in this context. Therefore, it is important to highlight the different nature of standards proxies that we use in this Chapter and in Chapter 3. In the latter, standards are applied in terms of mandatory and legal minimum requirements of food

TABLE 5.2: The effect of IFS on trade at the aggregate level: Results of ppml and IV estimations.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	ppml - no IV	IV (cfa)	IV (cfa, 1^{st})	ρ_2	IV (gmm)
IFS_lag	0.568***				
	(0.014)				
lnGDP_Importer	0.731***	0.836***	-0.023*		0.757***
	(0.023)	(0.053)	(0.014)		(0.054)
lnGDP_Exporter	0.385***	0.607***	0.435***		0.343***
	(0.014)	(0.041)	(0.007)		(0.042)
lnDist	-0.861***	-0.798***	0.011		-0.851***
	(0.054)	(0.094)	(0.034)		(0.121)
lnTariff	-0.741***	0.213	-0.235***		-0.367**
	(0.149)	(0.412)	(0.074)		(0.169)
Comlang_ethno	0.735***	0.070	-0.028		0.788***
	(0.066)	(0.189)	(0.061)		(0.148)
Colony	0.208**	0.543**	0.002		0.218
	(0.095)	(0.236)	(0.065)		(0.217)
Contig	0.251***	0.867***	0.022		0.211
	(0.068)	(0.149)	(0.054)		(0.146)
RTA	1.400***	0.815***	0.027		1.363***
	(0.112)	(0.152)	(0.053)		(0.248)
D_Bakery	0.249***	0.080	-0.024		0.278**
	(0.059)	(0.141)	(0.066)		(0.130)
D_Beverages	0.554***	0.741***	-0.506***		0.617***
	(0.063)	(0.153)	(0.064)		(0.141)
D_Dairy	1.006***	0.965***	-0.135**		1.024***
	(0.067)	(0.167)	(0.069)		(0.152)
D_Egg	0.530***	0.626**	-0.599***		0.646*
	(0.151)	(0.276)	(0.085)		(0.341)
D_FV	-0.673***	-0.575***	-0.081		-0.697***
	(0.060)	(0.150)	(0.065)		(0.132)
D_Fish	0.371***	0.945***	0.092		0.437***
	(0.059)	(0.165)	(0.063)		(0.142)
IFS		0.518***			0.635***
		(0.052)			(0.051)
IFS_Neighbour			0.443***		
			(0.005)		
Constant	-28.450***	-37.356***	-10.242***	0.097*	-28.319***
	(0.761)	(1.666)	(0.434)	(0.058)	(1.843)
Observations	59,458	56,043	56,043	56,043	56,043
R-squared	0.230				

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Baier-Bergstrand method used to account for multilateral resistance.
FV stands for Fruits and Vegetables.

TABLE 5.3: IFS estimation results by income group using PPML without and IFS-certification in neighbouring countries as IV

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	High Income				**Upper Middle Income**			**Low Income**			
VARIABLES	ppml - no IV	IV (cfa)	IV (cfa, 1st)	IV (gmm)	ppml - no IV	IV (cfa)	IV (cfa, 1st)	ppml - no IV	IV (cfa)	IV (cfa, 1st)	IV (gmm)
IFS_lag	0.513***				0.242***			0.445***			
	(0.015)				(0.076)			(0.038)			
lnGDP_Importer	0.726***	0.814***	-0.032*	0.766***	0.680***	1.186***	-0.001	0.788***	0.713***	0.003	1.050***
	(0.026)	(0.049)	(0.017)	(0.058)	(0.088)	(0.164)	(0.027)	(0.053)	(0.182)	(0.030)	(0.343)
lnGDP_Exporter	0.417***	0.671***	0.627***	0.384***	0.405***	0.577***	0.063***	0.219***	1.080**	0.289***	1.184
	(0.015)	(0.044)	(0.010)	(0.048)	(0.064)	(0.094)	(0.014)	(0.033)	(0.465)	(0.010)	(0.736)
lnDist	-0.835***	-0.737***	0.014	-0.834***	-0.627***	-1.942**	0.001	-1.317***	-6.045***	0.007	-5.087**
	(0.053)	(0.090)	(0.031)	(0.120)	(0.241)	(0.805)	(0.103)	(0.366)	(1.544)	(0.187)	(2.003)
lnTariff	-1.139***	-1.129***	-0.207**	-0.754***	2.420***	0.172	-0.097	0.075	-3.760	-2.181***	-8.415***
	(0.193)	(0.359)	(0.082)	(0.181)	(0.707)	(0.843)	(0.111)	(0.422)	(3.816)	(0.216)	(1.931)
Comlang_ethno	0.741***	0.078	-0.040	0.828***	-0.365	0.973	0.011	-0.286	0.238	-0.004	-3.348
	(0.066)	(0.154)	(0.057)	(0.146)	(1.538)	(2.518)	(0.328)	(0.359)	(0.638)	(0.103)	(4.465)
Colony	0.091	0.412*	-0.009	0.175	1.143***	0.127	-0.002	0.453***	0.389	0.006	1.926
	(0.103)	(0.230)	(0.078)	(0.233)	(0.341)	(1.031)	(0.120)	(0.158)	(0.503)	(0.141)	(1.175)
Contig	0.241***	0.827***	0.031	0.199	3.899	8.345*	-0.014	7.189***	8.640	0.043	7.431
	(0.068)	(0.138)	(0.050)	(0.146)	(3.200)	(4.618)	(0.825)	(2.479)	(5.356)	(0.949)	(8.061)
RTA	1.369***	0.775***	0.043	1.362***	-1.132***	1.691**	0.000	-3.339***	-5.830***	-0.607**	-6.200***
	(0.114)	(0.174)	(0.054)	(0.253)	(0.199)	(0.799)	(0.116)	(0.574)	(1.614)	(0.260)	(2.291)
D_Bakery	-0.121	0.066	0.144**	0.334**	-0.705*	-0.657	-0.069	-1.407***	0.522	-0.482***	-2.470
	(0.148)	(0.146)	(0.065)	(0.132)	(0.399)	(0.517)	(0.129)	(0.220)	(0.943)	(0.142)	(2.306)
D_Beverage	0.112	0.491***	-0.413***	0.602***	1.980***	-0.723	-0.665***	-1.646***	-0.604	-1.009***	-3.717
	(0.149)	(0.155)	(0.065)	(0.145)	(0.392)	(0.555)	(0.112)	(0.259)	(1.648)	(0.138)	(2.446)
D_Dairy	0.571***	0.743***	-0.071	1.036***	-0.298	-0.204	-0.544***	-7.797***	-10.044***	-1.317***	-11.956***
	(0.150)	(0.166)	(0.067)	(0.154)	(0.636)	(0.629)	(0.118)	(0.453)	(2.238)	(0.170)	(3.346)
D_Egg		0.323	-0.795***	0.597*		-2.417***	0.238**				
		(0.263)	(0.086)	(0.347)		(0.592)	(0.113)				
D_FV	-1.045***	-0.774***	-0.178**	-0.640***	0.277	-1.285***	-0.285**	-0.967***	3.007***	0.475***	0.166
	(0.150)	(0.154)	(0.070)	(0.137)	(0.407)	(0.477)	(0.116)	(0.227)	(1.106)	(0.139)	(1.332)
D_Fish	-0.176	0.308*	-0.132**	0.327**	1.756***	0.607	-0.019	0.627***	5.165***	0.547***	2.766***
	(0.148)	(0.169)	(0.066)	(0.155)	(0.390)	(0.439)	(0.112)	(0.229)	(1.169)	(0.140)	(0.644)
IFS		0.541***		0.603***		-1.649***			-1.564		-2.225
		(0.035)		(0.057)		(0.419)			(1.613)		(1.995)
IFS_Neighbour			0.508***				0.192***			0.034***	
			(0.006)				(0.014)			(0.010)	
Constant	-28.459***	-38.201***	-15.225***	-29.574***	-28.586***	-43.890***	-0.780	-25.283***	-46.472***	-6.547***	-56.844**
	(0.821)	(1.920)	(0.559)	(2.028)	(2.251)	(5.075)	(0.827)	(1.663)	(11.491)	(0.874)	(26.475)
Observations	32,382	34,721	34,721	34,721	14,329	13,272	13,272	12,747	8,050	8,050	8,050
R-squared	0.221				0.029			0.111			

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Countries of three income-groups are listed in the Appendix D in Table D.2.
Baier-Bergstrand method used to account for multilateral resistance.
FV stands for Fruits and Vegetables.

products. Negative effects on trade are straightforward to explain. In this Chapter however, standards are implemented as the number of certified firms in the exporting country. There is no immediate argument why more certified firms should reduce exports. However, it is possible that more IFS certification of fish farms crowds out exports of non-IFS certified fish such that exports of IFS-certified fish increase, but the net effect of fish exports is negative. This channel remains difficult to prove because it would require firm level data of (fish-) farms. Assuming that this explanation is correct, the effect on overall welfare remains unclear nevertheless. If the effect on total exports of a particular product remains negative even in the long-run, it is questionable whether the found positive effects in terms of better vertical integration (Colen *et al.*, 2012) offset lower exports.

IFS and the integration of developing countries into global export markets

The effect of IFS-certification on exports differs by income-level of the exporting country. If our hypothesis that high income countries benefit more from IFS certification in terms of increasing exports than countries with lower income-levels holds, the IFS coefficients should either decline in magnitude or become even negative. While the coefficient for high-income countries is as expected, it turns negative for upper-middle income countries and non-distinguishable from zero for low-income countries. These diverse findings require a detailed discussion.

First, we interpret our hypothesis as confirmed based on the heterogeneous findings which differ by income. Second, only high-income countries benefit from IFS-certification in terms of increasing export volumes. The negative effect for upper-middle income countries might be a result of similar mechanisms as explained for the negative coefficients on fish exports. Thus, the net effect remains unclear. Maybe IFS-certified products increases nevertheless, but the net effect remains negative. The non-significance for low-income countries might be due to the same effect only that the net effect is non-distinguishable from zero. Alternatively, findings indicate that IFS-certification is not sufficient to integrate low-income countries into global value chains. In other words, the trade-enhancing factors such as lower information asymmetries and lower transaction costs are not sufficiently relevant to offset factors of the trade-barrier view such as compliance costs and non-monetary barriers like weak public institutions (see Chapter 4). As shown in Figures 5.1 and 5.2, low income countries have on average lower compliance levels than high income countries. However, the non-significance of the IFS coefficient does allow us to conclude only that IFS certification in low-income countries does not affect exports. Hence, we do not know whether compliance costs are actually a relevant trade barrier in this context because the empirical model does not address the likelihood of compliance, but rather that compliance does not affect trade. This might be the case because the assumed reduction of information asymmetries is not sufficient to increase demand of importing countries. Consequently, even if compliance takes place, our findings do not provide empirical support for the conclusion that compliance as such is sufficient to increase country i's exports.

The latter interpretation might contradict findings of previous studies, especially about vegetables production in Senegal (Maertens *et al.*, 2009; Colen *et al.*, 2012;

Swinnen *et al.*, 2015). The authors emphasise – based on a rich firm- and household-level dataset – the potential of standards to integrate low income countries into global value chains, increasing rural income, and agricultural exports. However, although these analyses address multiple outcomes and not only trade, these studies remain case-specific such that external validity remains unclear. Moreover, most studies related to the Senegal-project are based on the private standard GlobalGAP (or formerly EuroGAP) which is a pre-farm gate and mostly addresses agricultural raw products. In contrast, IFS is a post-farm gate standard for processed food products. The effect on trade of such a post-farm gate standard might be less pronounced, because high-income countries have on average higher tariffs on processed food products. Moreover, food processing generates on average higher value-added than the production of agricultural raw products. Most high-income countries prefer to generate the value-added themselves. This phenomenon is known as "tariff escalation" and a serious threat for integration of developing countries into global value chains (Akyol *et al.*, 2005). Thus, a possibly trade-enhancing effect of IFS-certification in developing countries might be offset by these protections trade policies of high-income countries which are distinct for processed food products.

The type of food standard and θ

As explained in the conceptual model in Chapter 2, the decision of firms to invest in quality upgrading in terms of stricter standards is *c.p.* determined by the parameter θ, which represents the "ease" to address preferences of consumers. In contrast to Chapter 3 in which standards are found to reduce trade without exception, IFS-certification increases trade at the aggregate level, for almost all products included in the study (only results of fish are mixed), and for high-income countries. Apparently, firms are more likely to invest in IFS certification compared to MRLs which were found to reduce trade.

Neither the theoretical nor the empirical model specifies θ to keep both models flexible. In the case of IFS, θ is relatively large because of the trade-enhancing effect. In other words, exporting countries with relatively high numbers of IFS certification have an advantage for exporting processed food products compared to those with less IFS certification. It seems that IFS is particularly capable to reduce information asymmetries and to signal high quality of processed agricultural products. Other standards, such as MRLs as a relevant public standard, are less capable to signal high quality or do not allow firms to increase the mark-up. Thus, consumers are not willing to pay higher prices for these products. This is different in the case of IFS certification. First, because it certifies processed food for which the value-added is higher than for agricultural raw products. Moreover, IFS might be more able to signal high quality. Both is important since these affect the WTP of consumers. For example, Balogh *et al.* (2016) analyse the WTP of Hungarian salami and find quality certification as an important determinant of consumers' WTP. The authors conclude that producers can charge a higher price premium in these cases. All in all, by combining our research results from the macro-trade perspective with firm-level case-studies from the marketing-perspective we conclude that this price premium is important for producers of processed food products whether they comply with standards or not.

To conclude, this analysis aims to lift the debate of food standards and trade towards a higher level by looking more specifically on the characteristics of the standard. Future research is required to complement these results of the country-perspective with those at the food-processing firm level in low-income countries. Eventually, to evaluate the increasing relevance of food standards comprehensively, various perspectives ranging from macro-trade, marketing towards rural development studies, and many more need to be taken into account. Moreover, to further validate the argument of quality upgrading and WTP, future research should deal with effects on trade of B2C standards such as Fairtrade or the MSC. If the quality-upgrading argument in combination with higher WTP are a reasonable explanation, the trade-enhancing effect should be even stronger for these certification schemes.

6 The Hartz reforms and the German labour force

Abstract:

We estimate the impact of the Hartz Reforms, the most prominent labour reforms in Germany since the Second World War, on the German labour market. We adopt a cross-country program evaluation approach where, employing Synthetic Control Method (SCM) for comparative case studies, we utilise the characteristics of OECD countries to construct a counterfactual for Germany. Existing research primarily focused on the unemployment rate. We find that while the impact of the reforms on unemployment rate is weak – which is consistent with existing research – they increased labour force participation, particularly among women and the elderly. These findings are further supported by the finding that employment to working age population climbed as well.

This chapter is joint work with Devesh Roy and Abdul Munasib. Malte Ehrich's contributions are: Development of research question, conceptual framework, research design, data management, and writing of sections about political and legal background. Devesh Roy is the corresponding author and responsible for writing overall. Abdul Munasib is responsible for empirical estimations.

6.1 Introduction

While most of Europe including sizeable economies such as Italy, Spain, and Great Britain struggled following the 2008 crisis the German labour market remained strong and resilient. Despite a decline of 4.7% in GDP in 2009 and an output decline of over 18% in manufacturing, employment level in Germany remained at 40 million in both 2008 and 2009, and subsequently rose to 41.5 million (Rinne *et al.*, 2013). The same *Economist* magazine that had called Germany the sick man of Europe (Economist, 1999) called it a miracle in labour markets in 2010 (Economist, 2013).

Over the decade prior to 2005, unemployment rate in Germany stayed consistently above 8% and, in some years, reached double digits. The passage of the Hartz reforms, the most aggressive post-war labour reform in the country, began in 2003 and the final stage was instituted by 2005. Since 2005, having peaked at 11.3%, the German unemployment rate has been in a continuous decline in each subsequent year – to 7.5% in 2008 and ultimately to 5.5% in 2012. The Hartz reforms coinciding with the turnaround in the labour markets has naturally been of particular interest to researchers and policy analysts.

The flexibilities brought in by these reforms is often credited by some for the steep decline of German unemployment rate after 2005 and the limited impact of the downturn of 2008 on German economy (Rinne *et al.*, 2013). Overall, however, the literature finds a less clear connection between the Hartz reforms and unemployment rate in Germany: while some authors offer evidence of at least some positive impact of the reforms on unemployment rate (Krause *et al.*, 2012; Krebs *et al.*, 2013; Launov *et al.*, 2013) others strongly disagree (Dustmann *et al.*, 2014; Akyol *et al.*, 2013).

In principle, the Hartz reforms did aim to bring down unemployment rate quickly and reduce the average length of unemployment by a third. More importantly, however, the Hartz reforms aimed at encouraging labour force participation, notably for women and older persons, providing incentives for the unemployed to accept a job and strengthening job-search activities (Bouvard *et al.*, 2013). These reforms rolled out a series of laws that overhauled the unemployment benefit system and the organisation of labour market services.

Some of the specific measures included shortening of the period of entitlement to unemployment benefit, ending of options for early retirement, and reduction of employer social security contributions (Gaskarth, 2014). No other EU country has implemented such extensive reforms in all these areas in such a short time (Knuth, 2014). The Hartz reforms, therefore, are also presumed to have changed the nature of job creation itself in Germany (Spermann, 2011).

In this paper, we conduct a comparative case study to estimate the impact of the Hartz reforms on the German labour market. We find that the reform did not have a robust causal impact on the overall unemployment rate, which supports some of the existing literature. Our main focus, however, is on labour force participation. We find that the Hartz reforms raised labour force participation (LFP) specifically

among women and relatively older in the workforce.

Our results are further supported by the additional findings of increased proportion of employed in the working age population. It is important to note that the Hartz reforms may not exhibit any impact on the unemployment rate if both employment and labour force participation increased. Proportion of working age population employed captures the impact of the Hartz reforms on employment more cleanly.

We disaggregate the labour force by gender and age groupings to capture the differential impact of Hartz by worker type. To the best of our knowledge, the causal impact of Hartz on such labour market outcome disaggregated by age and gender remains unexplored. This is the research gap that we try to bridge. Burda *et al.* (2016) in a rare study do present data to show the increase in LFP after Hartz but do not estimate its causal link with reforms. We carry out a variety of robustness checks and provide additional evidence in support of our findings.

Hartz reforms changed the value of LFP at the margin. Those who were most likely to be out of labour force or were near dropping out were impacted by Hartz to join, rejoin or delay dropping out, whatever the case may be. Importantly, the groups that were impacted significantly in terms of LFP, women and older workers, had an increase in their employment rates as well. We estimate a positive and significant causal impact of Hartz on LFP for women and for workers in the age group 55-64.

In 1997, the minimum age to receive 32 months of unemployment benefit – the maximum duration of unemployment benefits – was raised to 57 years (Gaskarth, 2014). In effect, unemployed citizens over the age of 57 continued to be able to stop their job search and withdraw from labour force and still receive unemployment benefits. Prior to Hartz, the labour force participation rate for elderly was expectedly low (Gaskarth, 2014).

Given the magnitude and the extent of the Hartz reforms and the large observable changes in the German labour market, a proper identification of the causal impacts of the policy change is of crucial importance. This estimation, however, faces a number of challenges. First, the Hartz reforms, a country-wide policy, is specific to Germany. With just a single treatment unit, accurate inference is difficult, perhaps impossible, in a clustering framework (Buchmueller *et al.*, 2011; Donald *et al.*, 2007). We, therefore, adopt a case-study approach and use the Synthetic Control Method (SCM), which is devised to address precisely these kinds of situations (Abadie *et al.*, 2010).

Specifically, we use the OECD countries to construct a counterfactual for Germany. In studying the impact of German reunification, Abadie *et al.* (2015) uses an SCM setup with OECD countries. Note that this cross-country context is essential due to time varying events. For example, consider the great recession of 2008, which makes pre-Hartz Germany a poor counterfactual for the post-Hartz Germany.

In impact evaluation, the reliability of an estimate critically hinges on the accuracy of the counterfactual (Abadie *et al.*, 2010). Because of time-varying factors, both

observed and unobservable, the pre-Hartz period is unlikely to be an accurate counterfactual for post-Hartz Germany. Also, while constructing counterfactuals using other countries as the control group, it is highly improbable that a single country would make for an adequate counterfactual for a treated Germany; it is also unlikely that all control units would play equally important roles as counterfactuals.

SCM provides a systematic way to choose comparison units: the counterfactual is the weighted average of the control countries, where pre-intervention matching across a range of characteristics over a long period of time generates the weights. SCM also naturally lends itself to permutations or randomization tests for inference (Bertrand *et al.*, 2004; Abadie *et al.*, 2010; Buchmueller *et al.*, 2011; Bohn *et al.*, 2014).

It is important that our analyses cover a 25-year time span (1991-2013). With twelve years of pre-Hartz reforms data, we are able to achieve good pre-intervention fits, i.e., generate a good counterfactual. Equally important is the long post-intervention data (ten years) that offers a major advantage for assessing the employment impacts. The long time period is crucial to account for changes in technology, firm responses to the policy change, and individuals' responses to the changed labour market conditions. Moreover, policy responses are also often staggered and might require time to unfold.

Such a long term view, however, has accompanying econometric issues; it is difficult to expect that unobservable factors that can potentially bias the estimates would remain invariant over the entire study period. Therefore, unlike the standard difference-in-difference model, the fact that SCM does not assume unobservables to be time-invariant (Abadie *et al.*, 2010) is a notable strength of our estimation strategy.

6.2 The Hartz reforms: Background Information

6.2.1 The Labour Market and Unemployment in Germany

In 1999, The Economist magazine had called Germany the sick man of Europe (Economist, 1999). In the period following the unification, Germany had started to stagnate and between 1994 and 2002, at 1.6% its growth was slower than the EU average (Gaskarth, 2014).[1] The slow growth was combined with high rates of unemployment exceeding 13 percent. Abadie *et al.* (2015) look at the impacts of 1990 reunification on West Germany's GDP using SCM and find a significant gap in German GDP due to the event. The German labour market was perceived as highly rigid and – among other factors like the inefficient education system – as one of the major reasons for Germany's weak economic performance (Schäfer, 2003; Ochel, 2005; OECD, 2006). The OECD even urged for "further institutional reform [i.e. in addition to the Hartz reforms] of the Public Employment Service to better activate the

[1] Unification in Germany increased the labour force by one third, a large share of which was inadequately trained for immediate employment in an open market economy. Despite the need to first retrain the labour force and reshape the formerly centrally planned economy it was, however, also a political objective to adjust East German wages to the comparatively high West German levels as quickly as possible (Jacobi *et al.*, 2006).

unemployed" (OECD, 2006, p.8).

With surging unemployment, then German Chancellor Schröder convened the Hartz Commission in 2002 to reform the labour market. Chancellor Schröder and his cabinet comprising members of the Social-democratic party (SPD) and the Green party agreed on the so-called Agenda 2010 with various packages being passed by the parliament during the years 2003 to 2005. While Agenda 2010 reform addressed several sectors of the German economy, the labour market component called the Hartz reforms aimed primarily at reducing unemployment payments and the duration of eligibility to receive unemployment benefits, easing of hiring for temporary employment, and, arguably, changing the incentive structure for the unemployed to search for and accept a job.

6.2.2 The Agenda 2010: Structural change of German labour market policy

The Hartz Commission's recommendations comprised the following four crucial reforms (Gaskarth, 2014): (i) the creation of Personal-Service-Agency (PSA) to act as temp agencies to place unemployed people with employers; (ii) a grant for entrepreneurs, the "Ich-AG" (Me, Inc.), to encourage new businesses; (iii) benefit cuts of up to 30% if a person on unemployment benefits refused to take up a reasonable offer of work (the reforms made it a requirement to take up "reasonable" offers of employment or provide a valid reason for declining them thus shifting the burden of proof to the unemployed) (Gaskarth, 2014); (iv) merging social welfare benefits with long-term unemployment benefits.

The first set of reforms, Hartz I and Hartz II, were introduced on January 1, 2003. Hartz III addresses the structure of the Federal Labour Office and not labour markets directly. Hartz IV combines the former unemployment assistance and social security payments to a fixed and means-tested unemployment assistance (ALG II) (Nagl *et al.*, 2014).[2] The second part of Hartz IV deals with the duration of unemployment benefits (Arbeitslosengeld, ALG I).[3]

The standard entitlement period for unemployment benefit was one year. Afterwards, or if not eligible at all, the unemployed received ALG II. ALG II accounted for a maximum of 57% of the former net wage. Further, the social security assistance was aimed

[2]The previous welfare system in Germany was designed to maintain the unemployed in their current station until they could find a job they desired, and which matched their qualifications and experience. Benefit levels were high and were of long duration. The FLO gave a low priority to job search assistance and monitoring. Sanctions for failing to meet job search requirements were rarely applied. Linking unemployment benefits to the former salary set a high *de facto* minimum wage, meaning there was little incentive for the unemployed to find work (Gaskarth, 2014).

[3]Traditionally, the unemployment insurance in Germany has been based on three pillars: unemployment benefits (ALG I), unemployment assistance (ALG II), and social security assistance (Sozialhilfe). ALG I account for 57% of the previous net wage if the person had worked for at least 360 days in the preceding two years (Nagl *et al.*, 2014). With Hartz IV reforms, in addition to the switch from wage-dependent compensation to fixed-amount payments in ALG II, the means test was tightened.

at guaranteed access to basic human needs by a fixed payment.[4]

With Hartz, the entitlement period to receive ALG I reduced from 32 months to 18 months for workers aged 55 and over on February 1, 2006 (Ebbinghaus *et al.*, 2006). PSAs in effect aimed at placing the unemployed who found it hardest to find work, e.g., the young who lacked work experience or the long-term unemployed who may have fallen out of the habit of work. Hartz II reforms also initiated new types of employment, "Minijob" and "Midijob" which were short term and part-time employments with higher thresholds for taxes and social insurance payments for employees, and with less worker protection. Mini jobs were attractive for people who already had insurance from another source, e.g., through their parents or spouse, or as a pensioner. The employer pays a flat rate tax, which means that earnings from mini jobs do not have to be included in the household income tax assessment and are, therefore, not affected by tax progression (Knuth, 2014).

The unlimited duration of unemployment benefit payments used to be an extraordinary feature of the German labour market leading to replacement rates for the long term unemployed which were higher than in any other OECD country (OECD, 2006). Replacement rates of short-term unemployed, in contrast, were – and still are – comparable to many other OECD countries (Jacobi *et al.*, 2006).

6.2.3 The Unemployment Rate and the Existing Literature

Some existing literature studying the consequences of the Hartz reforms use calibration based on search models to simulate the impacts of Hartz reforms on unemployment (Krause *et al.*, 2012; Krebs *et al.*, 2013). These studies find that Hartz reforms significantly reduced unemployment in Germany. Launov *et al.* (2013) using structural modeling show that some measures in the Hartz package had no noteworthy impact on the unemployment rate in Germany whereas some other measures seemed to have played a role in reducing unemployment rate to a moderate extent.

Dustmann *et al.* (2014) question attributing German labour market success to Hartz reforms; according to these authors, the threat of off-shoring jobs to recently opened-up central and eastern European countries in the early 1990s and decentralised nature of employer-union negotiations allowed German firms to successfully push for limited wage growth, thus allowing them to mitigate the impact of recession. Also, the German recovery of the 2000s was due more to other economic factors, such as the sharp increase in demand from emerging nations (e.g. China) for capital goods.

Similarly, Akyol *et al.* (2013), based on micro evaluations, argue that various Hartz related measures do not show effects on variables that can be related to employment.

[4]The maximum duration of unemployment assistance was raised to 32 months for 57 years old in 1997. However, unemployed over the age of 57 continued to be able to stop their job search and withdraw from official unemployment status and still receive unemployment benefits. The elderly used these benefits as a "popular bridge between the exit out of regular employment and the entry into old age pension. Unemployment incidence was high among older workers with long tenure, and their labour force participation rate was very low." Unemployment rates for the elderly rose to between 20% and 25% in the mid-1990s (Gaskarth, 2014). Hartz fundamentally changed this system.

Ac-cording to Akyol *et al.* (2013) it is likely that employment increased due to a process of wage moderation that had already begun in the 1990s, which is an argument also put forward by Dustmann *et al.* (2014). Note that results that do not attribute reduction in unemployment to Hartz reforms are compatible with increases in LFP and rising employment rates for some segments of the actual or potential work force.

6.3 Estimation technique

6.3.1 A Case Study Approach with Synthetic Control Method (SCM)

There are a number of advantages to using SCM for this study. As discussed above, in creating a comparison group, neither every other country nor a single country without a comparable reform would likely approximate the most relevant characteristics of Germany where labour market reform was adopted. SCM, in contrast, provides a comparison unit (or synthetic) that is a combination of the control countries – a data-driven procedure that calculates "optimal" weights to be assigned to each country in the control group based on pre-intervention characteristics – thus making explicit the relative contribution of each control unit to the counterfactual of interest (Abadie *et al.*, 2003; Abadie *et al.*, 2010). In SCM, the researcher is forced to demonstrate the affinities between the affected and unaffected units using observed characteristics (Abadie *et al.*, 2010; Abadie *et al.*, 2015).[5]

Secondly, when aggregate data are employed (as is the case here) the uncertainty remains about the ability of the control group to reproduce the counterfactual outcome that the affected unit would have exhibited in the absence of the intervention. This type of uncertainty is not reflected by the standard errors constructed with traditional inferential techniques for comparative case studies. As Buchmueller *et al.* (2011) explain, in a "clustering" framework, inference is based on asymptotic assumptions that do not apply in our case as the focus is on one country Germany with its labour reform in 2003.

The comparison of a single country, Germany, against all other countries in the control group collapses the degrees of freedom and results in much larger sample variance compared to the one typically obtained under conventional asymptotic framework. The latter can seriously overstate significance of the treatment (Donald *et al.*, 2007; Buchmueller *et al.*, 2011). We, therefore, apply the permutations or randomization test (Bertrand *et al.*, 2004; Abadie *et al.*, 2010; Buchmueller *et al.*, 2011; Bohn *et al.*, 2014) that SCM readily provides.

Additionally, Abadie *et al.* (2010) argue that unlike the traditional regression-based difference-in-difference model that restricts the effects of the unobservable confounders to be time-invariant so that they can be eliminated by taking time differences, SCM

[5]Neumark *et al.* (2014), in the context of the impact of minimum wage legislations, point out that in several studies that adopted regression-based models, there were underlying assumptions of similarities across states (for example, categorization by region). Unlike the *ad hoc* strategies with a presumption of affinity, SCM demonstrates affinities of the donour pool states with the exposed state.

allows the effects of such unobservables to vary with time. In particular, Abadie *et al.* (2010) show that with a long pre-intervention matching on outcomes and characteristics, a synthetic control also matches on time-varying unobservables.[6]

Finally, because the construction of a synthetic control does not require access to post-intervention outcomes, SCM allows us to decide on a study design without knowing its bearing on the findings (Abadie *et al.*, 2010). The ability to make decisions on research design while remaining agnostic about how each particular decision affects the conclusions of the study is a safeguard against actions motivated by a "desired" finding (Rubin, 2001).

6.3.2 The Synthetic Control for Germany

A typical SCM analysis is feasible when one or more countries exposed to an intervention can be compared to other countries that unlike Germany were not exposed to the same intervention. In this paper, the outcomes are the different measures of employment/unemployment, the exposed country is Germany, the intervention is the set of the Hartz reforms that started in 2003, and the donour pool (unexposed/control countries) comprises the OECD countries.

Seeleib-Kaiser *et al.* (2007) argue that chancellor Schröder's social democratic government was not expected to introduce such wide-ranging cuts in the social insurance system. The fact that these reforms were largely unanticipated is useful from the point of view of program evaluation (Abadie *et al.*, 2010).

We construct a synthetic Germany (the counterfactual) using the "optimal" weights and then conduct placebo studies (or permutations/randomization tests) that help infer the statistical significance of the estimated impact of the Hartz reforms in the form of a post-intervention gap between the actual and the synthetic Germany. Details are in Appendix of this chapter.

6.4 Data

All data for the outcome variables are collected beginning with 1991 – establishing a long period prior to the enactment of the Hartz reforms in 2003 – and ending in 2015, the last period available for the outcome variables at the time of this study. All the outcome variables are taken from the OECD database.

The working age population refers to people aged 15 to 64. The labour force consists of individuals of working age who are available for work and have taken specific steps to find work. The labour force participation rate is calculated as the labour force divided by the total working-age population. The unemployment rate is defined as the number of individuals who are unemployed as a percentage of the labour force

[6] As Abadie *et al.* (2015) state, "only units that are alike in both observed and unobserved determinants of the outcome variable as well as in the effect of those determinants on the outcome variable should produce similar trajectories of the outcome variable over extended periods of time."

and it is seasonally adjusted. Since the labour force includes the total number of unemployed individuals plus those in civilian employment, the unemployment rate and labour force participation rate, in conjunction with the count of working age population, allow the construction of the measure proportion of working age population that is employed. We express this measure as a percentage.

The use of the OECD database, therefore, is of particular importance that presents the harmonized unemployment rate as comparable measure across countries (OECD, 2016a). While the universe of our donour pool is the set of OECD countries, due to missing values, not all countries can be in the donour pool for every outcome (Table E.1 has a list of donour pool countries).

An extensive set of predictors (i.e., variables for pre-intervention matching) was used to construct the synthetic control for Germany. First, variables were included to reflect the state of the economy and its structure: per capita income and growth in per capita income, consumption, public expenditure and investment shares of GDP, and an openness measure (from Penn World Tables, PWT 7.0-7.1).

Demographic variables such as population density, share of urban population, and infant mortality rates were also used as predictors; these were obtained from World Development Indicators (WDI) of the World Bank. To measure the state of the financial market, we draw from the WDI, domestic credit to the private sector and domestic credit by the banking sector (as percentages of GDP). Since education plays an important role in determining labour market outcomes, the average years of schooling in a country is also used as a predictor (Barro-Lee database). Finally, to capture the demographics component of the Hartz reforms, we include as a predictor the share of population aged 65 and over (from World Bank).

Table E.2 shows that, overall, the averages of the outcome variables for Germany were comparable to the donour pool countries over the study period. In terms of pre-intervention characteristics, the donour pool is very similar to the averages of all OECD countries. The same averages for Germany, however, are quite different from the donour pool, especially for population weighted distance, openness measure, population density, percentage of urban population in the largest city, infant mortality rate and trade union density. These differences indicate that a simple average of the neither the OECD countries nor the donour pool is likely to be a good counterfactual for Germany, and that a reasonable counterfactual would require a systematic weighting procedure as done under SCM.

6.5 Results

Separate SCM estimations were carried out for each of the outcome variables (the top panel in Table E.2 lists the outcome variables).

6.5.1 Main Estimates

We start our SCM estimation with labour force participation rate. The top two pictures in Figure E.1 show the impact of the Hartz reforms on labour force participation rate. The picture on the left shows a close fit between the actual and the synthetic prior to the intervention and a wide gap afterwards. The post-intervention gap is the measure of the causal impact of the Hartz reforms on labour force participation rate. As a next step, the statistical significance of the gap needs to be determined.

The permutations/randomisation test answers the question, 'How often would we obtain a gap as large as this if we had chosen a country at random instead of Germany?' This is the question that the placebo tests address. We therefore apply the SCM to each country in the donour pool (the placebos). The pictures on the right panel of Figure E.1 present results of the permutation/randomisation/placebo tests as described in Appendix E.1 (Abadie *et al.* (2010); Bertrand *et al.* (2004); Bohn *et al.* (2014); Munasib *et al.* (2015)). The gap between the actual and the synthetic for Germany is represented in the darker line and for every donour pool country in lighter lines. We see that Germany stands out from the placebo estimates (where the intervention did not happen). We thus infer that the impact of the Hartz reforms on labour force participation rate was significant.

Column 1 of Table E.3 presents the details of the pictorial presentation in Figure E.1 as described above. Pre-intervention absolute prediction error to mean ratio (APEMR) and pre-intervention Root Mean Square Prediction Error (RMSPE) are reported that describe the goodness of the pre-estimation fits.

The finding of the permutations or randomisation tests is described in terms of the following (see Appendix E.1.2 for details). The p-value of the post- to pre RMSPE ratio is very small (reported as 0.00), Germany has the top post-to pre RMSPE ratio rank (i.e., ranked 1), and the donour probability – the probability of obtaining a post-to pre RMSPE ratio as large as Germany's if one were to assign the intervention at random in the data – is also very small (4%). We, therefore, conclude that the impact of the Hartz reforms on labour force participation rate in Germany was statistically significant.

Column 1 of panel B in Table E.3 reports the "optimal" weights generated by the SCM procedure. We find that Mexico, United States, Switzerland, Canada, Italy, Sweden, and Portugal (in order) show up with the largest weights. Column 1 and 4 of panel B in Table E.3 reports a comparison of the pre-intervention characteristics between the actual and the synthetic Germany. We find them to be very closely matched. Repeating the exercise for the rest of the outcome variables we find that the following (the results in the tables and the figures are to be interpreted the same way as explained above):

TABLE 6.1: Summary of results

Outcome	SCM impact	Result reporting
Labour force participation rate	Increased	Column 1, Table E.3, Figure E.1
Labour force participation rate: men	No impact	Column 2, Table E.3, Figure E.1
Labour force participation rate: women	Increased	Column 3, Table E.3, Figure E.1
Employment in working age population (%)	Increased	Column 1, Table E.4, Figure E.2
Employment in working age population - men (%)	No impact	Column 2, Table E.4, Figure E.2
Employment in working age population - women (%)	Increased	Column 3, Table E.4, Figure E.2
Labour force participation rate: age 15-24	No impact	Column 1, Table E.5, Figure E.3
Labour force participation rate: age 25-54	No impact	Column 2, Table E.5, Figure E.3
Labour force participation rate: age 55-64	Increased	Column 3, Table E.5, Figure E.3
Unemployment rate	No impact	Column 1, Appendix C
Unemployment rate: men	No impact	Column 2, Appendix C
Unemployment rate: women	No impact	Column 3, Appendix C

In Table E.4 and Figure E.2, we present the SCM estimates of the impact of Hartz reforms on percentage of working population employed in Germany, again disaggregated by gender. Once again post/pre RMSPE ratio rank shows significant overall impact driven by women's employment rate. In most characteristics and in most outcomes there is not much difference between the synthetic and actual Germany. The exceptions are domestic credit to the private sector and to some extent education expenditure. Barring these exceptions, the synthetics closely replicate the treatment in eight out of ten characteristics that are compared.

6.5.2 Estimates of the impact of Hartz reforms on labour force participation by age groups

Next we look at the case of labour force participation by age groups viz. 15-24, 25-54 and 55-64 respectively. Figure E.3 presents the SCM estimate of the impact of the Hartz reforms on LFP of different age groups in the working population. The left panel in the figure shows a post-intervention gap, which is the estimated impact of Hartz reforms only for the work force comprising age group 55-64. Our placebo tests in the right panel shows that this gap is also statistically significant (Table E.5): p-value is small, post- to pre-intervention RMSPE rank is two, and the donour probability is very small as well (7%).

We also find that Greece, New Zealand, Turkey, and Luxembourg (in order) contribute the most to the construction of the synthetic Germany. And, we find close matches, in most pre-intervention characteristic, between actual and synthetic Germany (bottom panel C of Table E.5). Averaging over the post-Hartz period 2004-2013, we estimate that the reforms contributed to an increase in employment rate of women and workers in the age group 55-64. Importantly, the 2008-09 crisis did not alter the employment rates for the elderly. Note that the estimates show that even with recession, the supply side in terms of labour force participation for women and elderly continued to increase. Moreover the secular increase in employment rate continued unabated even during the crisis as an outcome because of Hartz reforms.

6.5.3 Robustness tests

In Table E.6 we present a robustness test where we re-estimate the impact of the Hartz reforms on all outcomes using a different set of predictors. In Table E.7 we report yet another robustness test: we go back to our main estimates in Tables E.3, E.4, and E.5 and this time exclude from the donour pool the country with the largest weight. This is a sensitivity analysis that tests how sensitive the estimates are to the presence of that particular country in the donour pool.

Since we find evidence of a statistically significant impact of the Hartz reforms on labour force participation of women and elderly and employment of working population particularly women, we conduct a series of tests to ensure that these estimates are robust. First, we expand the list of predictors. In Table E.6, results of estimation are presented in column 1. The two additional predictors that we now include are years of schooling and share of population 65 years or older.[7] Next, we use an alternative donour pool which has eight more countries including some of the newer entrants to the EU. Expanding the donour pool comes at a cost of a shorter pre-intervention time horizon (due to missing observations of the outcome variable for the first 6 years). In both these tests we find that the estimated impacts of the Hartz reforms remain statistically significant (in column 2, while the rank remains 1 for the earlier statistically significant estimated impacts, the p-value and the donour probability remain small).

It can be argued that the gap between the synthetic and the treatment units found so far is caused by the synthetic's inability to replicate the treatment's post-intervention outcome. To deal with this concern we use a "time placebo" test using a fictitious event of Hartz reforms as a falsification test (Mideksa (2013); Abadie *et al.* (2015)): We assign a placebo Hartz reforms intervention in Germany at some year earlier than 2003. Greater confidence can be attached to the results if the synthetic in this case replicates the treatment's outcome following this fictitious intervention (i.e., there is no statistically significant post-intervention gap).

In other words, we apply a placebo Hartz Reforms five years earlier (in 1998) and analyse post-intervention data (1999-2002). This placebo intervention shows no impact on outcomes like labour force participation and employment rates of specific groups.

6.6 Discussion

We estimate that the Hartz reforms did not causally impact unemployment in Germany. Is this finding compatible with markedly and steadily falling unemployment rate that coincided with the period after the Hartz reforms? Dustmann *et al.* (2014) in an influential paper point to significant reductions in the unemployment rate to the inherent flexibility of the German labour market that predates the Hartz reforms. Specifically, the authors argue that the specific governance structure of the German

[7] The same robustness check was also conducted for the outcome variables presented in Table E.3. The findings of Table E.3 were robust to this perturbation.

system of industrial relations made the labour market highly flexible. In particular, they point to the process that shifts setting of wages, hours, and other aspects of working conditions, from the industry- and region-wide level to the level of the single firm and even a single worker. Note that these features predate the Hartz reforms.

While we find that the Hartz reforms did not impact the overall unemployment rate (Table E.8)[8] we do find that they had a significant causal impact on labour force participation rates, specifically for women and the older workforce. Our findings are buttressed by the additional findings of increased proportion of employed in the working age population.

The push towards greater labour force participation would be consistent with the changes related to unemployment benefits. Recall that the maximum number of months of unemployment benefits was reduced from 26 to 12 months for those under 55 years of age, and from 32 to 18 months for those 55 and older. With this reduction, the transition to employment could certainly be hastened and, for that, temporary employment could be the outlet for many workers. The employment rate has risen steeply for people aged 55 to 64, from around 37% in 2000 to 61.5% in 2012. In 2012, there were more people in the 60-65 age group in employment than in retirement for the first time since 1974 (Federal Institute for Population Research, 14.08.2013).

The survey by Kettner *et al.* (2007) finds that unemployed persons increased their job seeking activity and made greater wage concessions during job interviews. As another indicator of temporary employment, among people assigned to the PSAs, under 25 year olds were significantly over-represented.

As part of the reforms, spouses or partners of benefit recipients were strongly incentivised to seek work which was accompanied by the phasing-out of early retirement options (Bouvard *et al.*, 2013). In Germany, economically active population increased by 4.9% between 2004 and 2011, despite a decline in the working age population, on the strength of the 4.6-point rise in the participation rate.

The increase in participation rate is particularly stark for workers near the retirement age from 43% in 2003 to 67% in 2013 (OECD, 2016c) As such, the increase in participation was particularly high among older persons (16.2 points) and women, for whom the upward trend gained momentum over the period, rising 6.0 points, while the male participation rate rose by 3.3 points (Bouvard *et al.*, 2013). These findings are in line with our finding of a significant impact of the Hartz reforms on labour force participation rate.

Burda *et al.* (2016) state that it is also possible that some workers in temporary employment could be doing so as an additional job. Burda *et al.* (2016) find that in 2014, 8% of part-time as well as full time employees had a second job besides their main occupation. The moonlighting effect along with mobilisation of inactive

[8]Note that the estimates in Table E.8 show that while the Hartz reforms had a significance effect on overall unemployment rate and women's unemployment rate show in Panel A, subsequently in Panels B and C, they turn out to be non-robust.

workers following the Hartz reforms could also be contributing to the rise in number of temporary employees.

Burda *et al.* (2016) argue that the strong German labour market performance since 2005 is entirely attributable to the extensive, rather than the intensive employment margin, i.e., the supply of workers that were employed at a given working-age population. In 2014, total aggregate hours worked by all persons in Germany was merely 0.4% higher than in 1993, while employment in persons rose from 37.8 to 42.7 million persons.

Hence there has been a marked reduction in average hours. Burda *et al.* (2016) findings are consistent with a positive shock to labour supply, in particular, in West Germany and thereby underscore the role of the Hartz reforms in increasing employment. Our findings pick up the causal impacts of the Hartz reforms on labour supply. Our results are also compatible with near stagnant work hours. Some part of this can be attributable to the Hartz reforms as it is shown to alter the composition of employment. While the employment rate hit a historic high of nearly 73% in 2012, the total number of hours worked per worker declined from 1564 in 1992 to 1366 in 2014 (OECD, 2016d). The outlet of temporary employment would likely have worked in favour of this trend.

This significant increase in labour force participation of the older workers in such a short time most likely came through temporary employment. Workers near retirement are unlikely to be offered permanent jobs (Gaskarth, 2014). Also as part of the reforms, the unemployment insurance benefits were accompanied by the abolition of a tax-free allowance up to €11,000 for severance pay. Since severance payments are linked to tenure (how long has the person worked), workers, particularly older ones, would prefer to increase the tenure of employment by taking up temporary jobs.

It is important to note that temporary employment constitutes a relatively smaller percentage of the overall labour force, varying between 10% and 15% in Germany (OECD, 2016b). Yet, amongst individuals entering and leaving the workforce, the share of temporary agency workers is considerable. According to Spermann (2011), while the number of temporary employment relationships that begun and ended was between 600,000 and 700,000 per annum in 2000–05, in 2006–08 both figures had grown to over one million.

Following the Hartz reforms, some companies also started restructuring their labour force to shift towards more temporary employment, a substitution effect. Spermann (2011) presents the example of in-house staffing agency formed by Schlecker, a large drug store chain. Dauser (2009) presents a number of other examples as well. Crimmann *et al.* (2009), however, argue that in house temporary staffing is not a widespread phenomenon.

The creation of public interest employment in 2005 in the form of "one-euro" jobs as part of Hartz IV reforms was also a step towards temporary employment and increasing labour force participation of some groups. There were 200,000 one euro jobs in 2005 and it rose to 320,000 in 2009. Billed as a pathway to a permanent job for those that were unemployed or not even looking, it could have also spurred their job

search. The maximum working time in these jobs was 30 hours. These jobs typically lasted between six and nine months.

The other important aspects of the Hartz reforms are likely to be mini-jobs and midi-jobs that refer to low-wage employment contracts for a small number of hours worked. Mini-jobs existed before the Hartz laws although Hartz raised the monthly salary cap from €325 to €400, and created a higher level, called midi-jobs, which were capped at €800.[9] Midi-jobs are subject to lower social insurance contributions, which gradually rise to the standard rate when monthly earnings reach €850. Workers in midi-jobs are entitled, overall, to the same benefits as employees subject to social insurance contributions at the full rate.[10]

The numbers working in mini-jobs on top of other paid work (who are already counted in the employment figures because of their main job) have grown sharply since the Hartz reforms: From 1.69 million in 2004 to 2.53 million in 2011; on the other hand, those working only in a mini-job increased by just 81,000 between 2004 and 2011 (to 4.9 million, 66 percent of whom are women). The number in midi-jobs rose from 1.19 million in 2007 to 1.37 million in 2011 (74 percent of whom are women).

6.7 Conclusions

Contesting the objective of European policy leaders of achieving economic convergence, the current situation in Europe may very well be characterised by divergence since a specific set of countries are performing systematically better than the rest of Europe. One such country is Germany, the largest country in Europe both in size (physical and economic) and population.

We investigate the impact of the Hartz reforms on overall German labour market. We opted for a method that fits the context of estimation of the impact of an aggregative policy with a single treatment unit. The need was also felt for a method that is able to address the role of observed and unobserved factors to the best extent possible. These motivated us to use the SCM in assessing the impacts of the Hartz reforms on the German labour market. Our method of SCM is quite rewarding in this context.

While the results show no significant impact of the Hartz reforms on unemployment rate, we find that labour force participation increases, especially among women and

[9]Mini-jobs are subject to employee social insurance contributions at a reduced or zero rate. Employer contributions for jobs earning less than €400/month are higher than for those for a 'regular' job, approximately 28% instead of the roughly 20% standard rate for employers. Workers in mini-jobs are not entitled to unemployment insurance benefit, and pension benefits are optional in that workers may opt to pay a pension contribution at a reduced 4.9% rate, which gives them pension rights similar to the standard contributions. A mini-job does not in itself entitle the worker to health insurance insofar as the employer does not pay health insurance contributions. The individual may nevertheless be entitled to health insurance if it is provided under another activity, or if covered through a family member.

[10]A workfare arrangement in the non-private sector established by Hartz IV, under which the claimant continues to receive benefits, in addition to payment of at least one euro an hour for work in the public interest.

the elderly. We also find an increase in the proportion of working age population employed.

Finally, what do the results imply for other countries? It has been suggested by some researchers that other countries, particularly the core countries, in Europe could benefit from their own Hartz type reforms (European Central Bank, 2012). This argument is conditioned by the belief that the Hartz reforms caused significant changes in unemployment in Germany, which is not supported by the estimates in this paper. Also, several labour market policies that characterise Germany such as no minimum wage (it was introduced only recently) might not be similar in other countries. This implies that the employment outcomes that we show as impacted by the spurt in labour force participation could be quite specific to Germany.

Our analyses point to opportunities for a great deal of future research. The pathway of rise in temporary employment and no change in unemployment is suggestive. It needs to be strengthened with longitudinal study that would have tracked different types of workers (firms) as well as job seekers (employee seekers) before and after the Hartz reforms. This would be important to explore the motivations (incentive structures) of the workers (firms).

Further, there is ample evidence in the literature that East Germany and West Germany behave differently in terms of their respective labour markets (Snower, D., Merkl, C. (2006); Burda (2006)). Exploring these differences will not only provide additional insight into the regional variations but may also offer valuable insights into the long-term effects of sociopolitical institutions.

7 Conclusion

This final chapter summarises the results of all four essays, discusses limitations, policy implications, and open questions that we could not address and therefore, devote to future research. Moreover, we relate the essays into the broader research context.

The raising relevance of food standards constitutes the point of departure of the first three essays of the thesis. Because standards are a subset of NTBs overall, their effect on international trade is a matter of concern. Theoretical as well as empirical results remain ambiguous. The entire debate has been entitled as "standards-as-catalyst vs. standards-as-barriers" to trade. All three essays aim to provide additional insights *why* the effect is neither positive nor negative in all settings. Instead, the thesis provides arguments under which conditions which effect predominates. Whether the effect is positive or negative depends on market structure, complexity of products, public institutions, the type of standard, and the income level of the exporting country. Consequently, we contribute to the debate by suggesting a compromise in the sense that both views are correct, depending on the research context.

7.1 Overview of results

The first essay (Chapter 3) focuses on the role of export volume as additional determinant for the effect of standards on trade. Based on the argument that compliance costs are primarily of fixed costs character, the hypothesis claims that leading exporting countries benefit from stricter standards in terms of trade volume because these are more likely to invest while country-pairs with relatively low trade volumes do not invest and trade less. Based on a quantile regression framework within the gravity model of trade, we indeed find empirical evidence of heterogeneous effects of standards and of tariffs on trade. As standards become stricter, country-pairs from the first to the eighth decile are expected to trade less while country-pairs at the highest decile trade more. The discussion addresses potential implications for the degree of concentration of global agricultural export markets.

The role of fixed costs of compliance with standards remains the point of departure of the second essay (Chapter 4) as well. We claim that exports to markets with relatively strict standards does not only require sufficient income-levels of the country of origin, but also non-monetary factors like public institutions need to be of sufficient quality. The essay shows that the quality of public institutions is especially relevant for exports of rather complex products like dairy- and meat-products. We use the EU as an example of a destination market of sufficient size and relatively strict standards. Thus, the quality of public institutions of the country of origin as well as the degree of product complexity are important factors as well which determine whether standards

increase or reduce trade.

The third essay (Chapter 5) is related to the previous essay by focusing on more complex agricultural products, which are processed foot items in this case. The number of IFS certified producers is the main variable of interest within a gravity model of trade. This is a particular interesting case because IFS data, as an important private food standard, are difficult to obtain. Moreover, IFS is a post-farm gate standards which refers to the upper part of the value chain in contrast to other private standards like GlobalGAP which are more frequently analysed. In contrast to the first essay in Chapter 3, standards are found to increase exports at the aggregate level. However, the effect remains robust only for high-income countries and disappears for other, that means lower income groups.

Overall, we contribute to the catalyst- and barriers-debate by highlighting important determinants which affect the direction of the coefficient of interest. Standards are likely to reduce trade between country-pairs which trade little in any case but increase trade between country-pairs with well-established trade flows. Moreover, the quality of public institutions is an additional determinant to what extent countries can meet standards of destination markets. Furthermore, product complexity matters. Finally, the type of standard is relevant, too. Arguments of the catalyst-perspective such as lower information asymmetries and reduced transaction costs seem to be more relevant for the private standards IFS certification whereas compliance costs offset potential trade-enhancing factors in the case of the public food safety standard MRL.

The final essay four of Chapter 6 is unrelated to the previous studies and analyses the effect of the most comprehensive labour market reform in Germany on various outcome variables like unemployment and LFP. Based on SCM, we find weak effects of the reforms, which aimed to make the labour market more flexible and dynamic, on unemployment. However, LFP increased due to the reforms especially of women and elderly people. This is an important contribution to the tensed and intensive public debate of successful labour market reforms in particular during an era of multiple crises within Europe. Our findings raise doubt about the often stated argument that unemployment has declined as a result of these reforms. In line with the few other existing previous studies, we do not find support that the reforms have caused the decline after 2005.

7.2 Limitations

All essays are subject to limitations to at least some extent. The rules of good scientific practice require to state those explicitly to make contributions to scientific debates of this thesis as transparent as possible. Moreover, stating limitations clearly facilitates future work to address those as adequate as possible.

Chapter 3 could have been improved by estimations at the product-level. Earlier versions of the chapter were based on analyses at the product-level. However, as the empirical strategy became more advanced to address all empirical challenges, we faced severe problems due to collinearity and non-convergence. MRLs are given

at the product level and it is reasonable to assume that effects on trade differ by product. Since we included country-pair-product fixed effects, we are still convinced that our estimates remain unbiased. However, an analysis at the sectoral level would have been interesting and could provide more precise policy implications. Moreover, the theoretical foundation of gravity suggests to use production and consumption as proxies for the economic size of trading partners. However, we faced significant data problems because production data at the disaggregated level remain often incomplete. This problem would have been amplified because consumption data would have been computed by simple accounting techniques as the sum of production and imports minus exports. Finally, other indices like Winchester *et al.* (2012) and Achterbosch *et al.* (2009), which are dyadic, could be considered as additional robustness checks. The interpretation of these two indices differs from the Ferro-index. The Ferro-index measures the restrictiveness of the importing country relative to all other countries. In other words, it indicates the multiregional restrictiveness of a particular country. In contrast, the other two indices of Achterbosch and Winchester are measures of the relative difference in terms of restrictiveness within a particular country-pair. Thus, the interpretation refers more to the "bridge-to-cross" (Munasib *et al.*, 2011) assuming that exporting countries that are less strict than the importing country needs to spend more effort in meeting the standard of the importing country than countries that have similar standards.

Strictly speaking, the empirical model in Chapter 4 does not completely account for multilateral resistance. It relies on the Baier-Bergstrand method, which is overall well accepted, but mostly at the aggregate level. Since we use five different product categories, this procedure is not ideal. If we were able to use fixed effects, we would have needed country-year-product fixed effects. This was computationally not feasible. However, we include product fixed effects instead to account for multilateral resistance as good as possible. In addition, instruments need to be considered as further robustness checks.

The limitation of the previous chapter regarding Baier-Bergstrand at the sectoral level is partly valid for the analysis in Chapter 5 as well. Moreover, the type of standard proxy is not ideal. Whereas we use requirements of importing countries in Chapter 3, this essay implements standards as the number of certified producers of the exporting country. However, we do not know how much each producer exports. We also do not know the share of certified exports of total exports of a particular product. In addition, to answer the research question precisely, we would need a measure which indicates the restrictiveness of importing countries with respect to IFS as we do in the MRL-setting. A negative estimated coefficient would then indicate support for the trade-barriers view. In contrast, a negative sign of IFS certified producers is less straightforward to interpret. But since most standards also refer to qualitative requirements, most will remain difficult to quantify. Another limitation might be that total number of IFS certification is an absolute measure. Instead, a relative variable would be more adequate that is for example weighted by the number of producing firms. However, we are not aware of any data source that provides number of firms at this disaggregated level.

The last essay about the German labour market and how it is affected by the Agenda 2010 relies on the argument that no other systematic treatment took place during

the years 2003 to 2005 which might bias our results. However, we believe that the arguments and robustness checks discussed in the chapter are comprehensive and convincing.

7.3 Policy implications and open discussion

Policy makers need to take into account the trade-off between potential trade-reducing effects as well as exclusion of small-scale farmers and preferences of consumers for food products that meet specific standards in various dimensions. Although there is no argument related to food safety why requirements should differ across human beings, preferences regarding food safety and other objectives like environmental aspects, animal welfare, and labour standards do differ across consumers and countries. This inhibits complete harmonisation of standards as policy recommendation because it might actually provide additional arguments for anti-trade advocates. If standards become harmonised for a given country, harmonised standards are either lower or higher on average than before. If these turn out to be lower, consumers' concerns about food safety might increase and therefore, anti-trade attitudes. If harmonised standards are stricter on average, trade might be even more adversely affected even though harmonisation should increase trade in principle.

Although a lower amount of varying standards might be desirable from a trade-policy point of view, governments should rather support farmers to meet standards by providing required infrastructure, credits, and knowledge. The raising relevance of standards should not result in additional advantage for leading exporting countries as it is currently the case (see Chapter 3). Instead, since standards do affect international trade, these should be least-distortive to trade as possible. This is more likely to be the case if compliance costs are low or covered by other sources.

Since the effect of standards on trade differs by the degree of complexity of a particular product, tariffs of processed food products should be further reduced to support developing countries to participate not only in value chains but also in high value chains. We do not find empirical evidence for the ability of private standards to enhance the integration of developing countries into global value chains. Consequently, we can not conclude that private standards constitute a one-suits-all development tool.

In terms of public standards that are covered by SPS-regulation, governments need to ensure scientific justification of standards to reduce the risk that public standards are implemented due to protectionist motives. This implies to strengthen the WTO as the major global institution which designs and enforces the regulatory framework for international trade. It remains questionable to what extent regional trade agreements are useful substitutes for a period during which the WTO and the Doha-round remains blocked by several member states. However, if countries manage to form supranational institutions like the EU which negotiates trade agreements and standards for their member countries, these at least partly harmonise standards within this particular group of countries.

As it became evident during negotiations of the Transatlantic Trade and Investment Partnership (TTIP), consumers are concerned about food safety standards. This prohibits complete harmonisation and an overall drastic reduction of standards. At the first glance, these concerns need to be addressed seriously and transparently independent of their scientific justification to avoid future raises of anti-trade movements. This implies to provide communication platforms which allow consumers to express their opinion and to discuss related arguments. Appearance, participation, and credibility of scientists of related disciplines are essential at this stage. It is not only the case that all parties involved often provide and believe in studies which favour their individual views, it is also the increasing non-acceptance of scientific evidence which is problematic as it became evident with the term *fake-news*. Participants of public debates seem to ignore arguments of their opponents which makes finding a compromise increasingly difficult. However, the willingness and empathy to compromise is essential not only in debates about free trade and regional trade agreements like TTIP but for the functionality of democracies overall. Consequently, it is about scientists to obey rules of good scientific practice and to discuss their arguments in public such that the general public and involved groups with varying interest trust these findings. In the case of standards and trade, this implies to state explicitly to what extent a specific residue is harmful at which levels of consumption. Finally, wrong arguments need to be addressed as well. For example in the case of TTIP, I have rarely observed scientist who argue against the rumour that the US has lower food standards on average than the EU.

The final policy recommendation of the three trade essays does not directly follow from the conducted analyses in this thesis but are crucial for the debates about standards and trade as well. The founders of the German constitution[1] had good reasons to limit power of plebiscites. Public opinions can be easily misguided by powerful groups with specific interests. It is the author's belief that this partly happens during TTIP-debates which were often not based on justified arguments. Since decisions on food safety and international trade are complex as well and many powerful interest groups are involved, these decisions are not binary and should not be made by plebiscites but only by representative parliaments.

Overall, we conclude that standards do affect international trade and are likely to affect the trade-GDP ratio. Our main argument is that standards can either enhance or reduce trade depending on various parameters. Therefore, the net effect of standards on the trade-GDP ratio remains unclear such that predictions regarding changes of the global slow-down of trade – as discussed in Chapter 1 – as a consequence of raising relevance of standards are difficult. But it is not difficult to conclude that analyses regarding the effect of standards on trade remain important. If nationalist movements intensify, governments might make use of hidden protectionist measures such as NTBs more frequently. At the same time, global institutions such as the WTO or the United Nations which are fundamental for the regulatory framework of global exchange need to be strengthened or at least sustained. The crucial role and responsibility of the EU – and hence Germany as its largest economy – can not be overemphasised in this respect.

[1] Strictly speaking, Germany does not have a constitution but only a *Grundgesetz* which was meant to be temporary. However, meaning, significance, and functioning is the same as for constitutions. But no English translation exists for the German-specific term *Grundgesetz*.

Policy implications of the fourth essay are manifold. First, the Agenda 2010 increased LFP of women and elderly people. Traditionally, these are groups, that are often less frequently available for the labour market. The increasing LFP of women is desirable in general. Whether or not elderly between 55 and 64 shall be part of the labour force remains a normative decision of each society. However, there are valid arguments in favour of increasing LFP of both groups. If these groups continue to work, they do not only contribute more and longer to economic output, an aging society also relies on an experienced and diverse workforce. Furthermore, it supports the public pension scheme which relies on a working population of sufficient size. Second, policy makers should not implement policies which are erroneously built upon pure correlation. Similarly, economists should not recommend policies based on pure correlation as done for example by (Goecke *et al.*, 2013). This discussion refers directly to the previous argument above about the role of economists and scientists in general. Providing policy recommendations and results of impact evaluation packed in a box which is labeled as science but which are *de-facto* based on non-scientific work undermines credibility of science and *real news*.

However, external validity of our results remains ambiguous. In contrast to Germany, youth unemployment is a major problem in many other European countries in particular in the south. The Agenda 2010 addressed primarily the supply side of the labour market, i.e. groups without employment. Increasing incentives by lowering unemployment benefits to accept job offers remains pointless if there aren't any or only very few jobs available. In these cases of severe shortage of employment opportunities, other policies are required which facilitate creation of jobs and improves the functionality of the economy in general.

A Appendix Chapter 1

A.1 Literature overview: notes for Table 1.1

TABLE A.1: Notes for the literature overview of Table 1.1.

a	Chad, Egypt, Gambia, Mali, Nigeria, Senegal, South Africa, Sudan, Zimbabwe
b	Austria, Belgium, Denmark, Finland, France, Germay, Greece, Ireland, Italy, Luxembourg The Netherlands, Portugal, Spain, Sweden, United Kingdom, Norway
c	STIC Reveision 2, "cereals and cereal preparations" and "fruits, nuts and vegetables"
d	SITC Revision 1, edible groundnuts (05172), groundnut oil (4214), shelled groundnut (2211)
e	The authors find a positive effect for developed countries and a negative effect for developing countries. In addition, leading exporters benefited from HACCP whereas small-scale producers faced a negative effect.
f	The number of included importing countries varies. The EU-15, USA, and Japan among others are always included.
g	The comtrade commodities are garlic (070320), onions (070310, 071220), and spinach (070970, 071030).
h	17 developing countries are included: Argentina, Bulgaria, Chile, Czech Republic, Honduras, India, Iran, Jordan, Kenya, Mozambique, Nigeria, Pakistan, Panama, Poland, Senegal, South Africa and Uganda.
i	EU, USA, Canada, Japan, Australia
j	EU27, Argentina, Brazil, Chile, China, Mexico, New Zealand, South Africa
k	Australia, Canada, Republic of Korea, Japan, Mexico, Russian Federation and USA
l	The direction of the effect highly depends on the corresponding country pair.
m	The corresponding commodity codes are (081330) for dried apples, (200970) for apple juice, and (200840) for pears.
n	The importing countries are: US, Mexico, Canada, Colombia, Brazil, Netherlands, United Kingdom, Spain Hong-Kong, China, Japan, Saudia Arabia, Arab Emirates, India, and Russia.
o	The authors esimated two models: the first applies an aggregated index of stringency whereas the second index is disaggreagted into its different dimensions, namely phytosanitary regulations, maximum residue limits, good agricultural practice (GAP), and quality standards.
p	Apart from the GAP-index which is found to have a positive impact, the disaggregated indexes are not statistically significant.
q	The exporting countries are: Argentina, Brazil, Chile, Colombia, Costa Rica, Ecuador, Guatemala, Honduras, Mexico, Panama, Jamaica, St. Lucia, Cameroon, Cote d'Ivoire, Morocco, South Africa, China, India, Indonesia, Philippines and Taiwan.
2	The importing countries are: Belgium, Luxembourg, France, Germany, the Netherlands, UK, Canada, Japan, New Zealand, Switzerland, and the US.
s	SITC Revision 1, code 0513.
t	EU (treated as one entity), Argentina, Australia, Brazil, Canada, China, Japan, New Zealand, Russia and the United States.
u	HS 6-digit commodities for fruits (080810 and 080820), vegetables (070110, 070190, 070200, 070930, and 070960) and cereals-grains (100300,100510, 100590, 120510, and 120590).
v	The authors calculated several heterogeneity indexes. All of them had no effect on trade flows, except MRL.
w	See Table B for a list of exporting and importing countries.
x	These are plant products ranging from HS-06 to HS-12.
y	The authors provide evidence for trade enhancing effect due to higher demand and, simulatenously, trade reducing effects due to reduced supply. This is interpreted as one reason for the mixed results in the corresponding empirical literature.
z	The exporting countries are: Kenya, Sri Lanka, China, India, Indonesia, Vietnam, Argentina, and Iran.
aa	The importing countries are: France, Germany, United Kingdom, Italy, Spain, Portugal, Austria, Ireland, Belgium, Denmark, Greece, Holland, Sweden, and Finland.

B Appendix Chapter 3

B.1 The partitioned iterative algorithm and high-dimensional fixed effects

As a point of departure, consider the normal linear regression model:

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon \tag{B.1}$$

The vector of estimated coefficients obtained via OLS is given by the following well-known formula:

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \tag{B.2}$$

However, depending on the field of research datasets might contains several million observations. Whenever controlling for group heterogeneity by including fixed effects becomes necessary, the amount of coefficients to be estimated increases as the sample-size increases. For example, the more countries, products, and years we include in our dataset, the more fixed effects are required be estimated. In other words, if the matrix $\mathbf{X}$ increases due to more fixed effects, computing the inverse of $\mathbf{X}'\mathbf{Y}$ might become infeasible due to computational constraints, especially due to limited memory capacities of computers.

Therefore, Guimaraes *et al.* (2010) propose a partitioned iterative algorithm to solve for $\hat{\beta}$ which is inefficient but has lower memory requirements. Thus, we decompose Equation (B.1) into a matrix $\mathbf{Z}$ which contains all explanatory variables, and a matrix $\mathbf{D}$ of dummy variables to account for fixed effects.

$$\mathbf{Y} = \mathbf{Z}\beta + \mathbf{D}\alpha + \epsilon \tag{B.3}$$

which can be re-written as

$$\begin{bmatrix} \mathbf{Z}'\mathbf{Z}\beta + \mathbf{Z}'\mathbf{D}\alpha = \mathbf{Z}'\mathbf{Y} \\ \mathbf{D}'\mathbf{Z}\beta + \mathbf{D}'\mathbf{D}\alpha = \mathbf{D}'\mathbf{Y} \end{bmatrix}$$

Solving for β and α yields:

$$\begin{bmatrix} \beta = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'(\mathbf{Y} - \mathbf{D}\alpha) \\ \alpha = (\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'(\mathbf{Y} - \mathbf{Z}\beta) \end{bmatrix} \tag{B.4}$$

Equation (B.4) shows that an estimate for β is obtained via a linear regression by only using $\mathbf{Z}$ and a transformed $\mathbf{Y}$. Ultimately, β and α are estimated according to the following procedure:

I $\mathbf{D}\alpha$ is added as an additional covariate to the regression of $\mathbf{Y}$ on $\mathbf{Z}$. The elements of $\mathbf{D}\alpha$ are initially equal to zero but are subsequently replaced until convergence is achieved.

II We save the sum of squared residuals of the regression of step one.

III A temporary variable is computed which is the difference between the dependent variable $\mathbf{Y}$ and the fitted values of the regression in step one plus the product of $\mathbf{D}\alpha$ and its coefficient of step one.

IV The elements of $\mathbf{D}\alpha$ are replaced by the group-average of the above mentioned temporary variable.

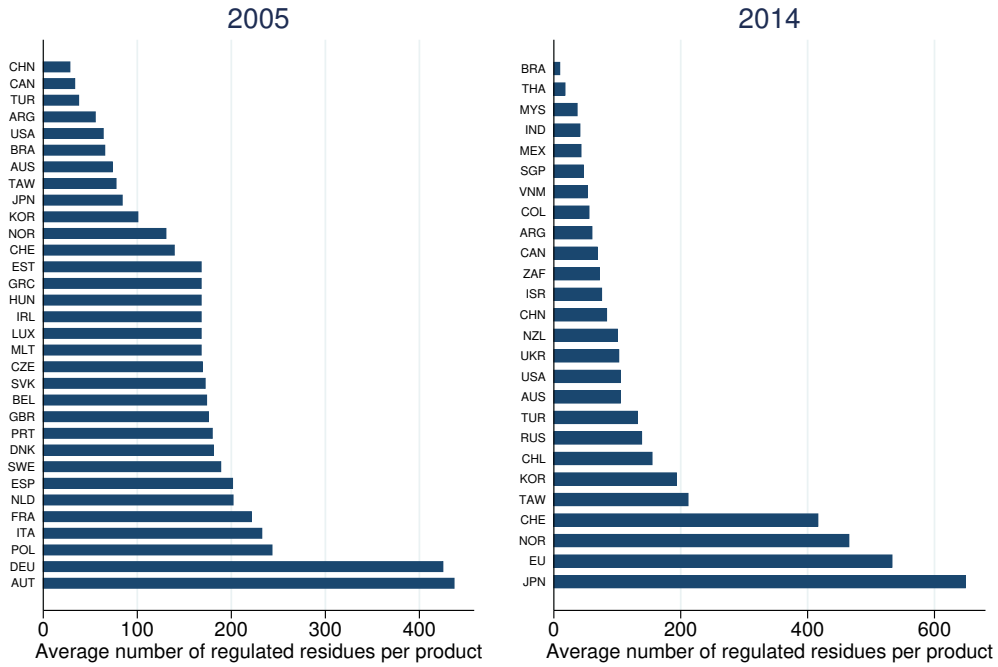
All steps are repeated until the coefficient on $\mathbf{D}\alpha$ is equal to one. The covariate $\mathbf{D}\alpha$ contains all estimated coefficients of the fixed effects which we need to subtract from the dependent variable, see Equation (B.4). Eventually, the new transformed dependent variable will be used in the quantile regression.

B.2 Descriptives on MRLs

TABLE B.1: Number of regulated residues per country and year

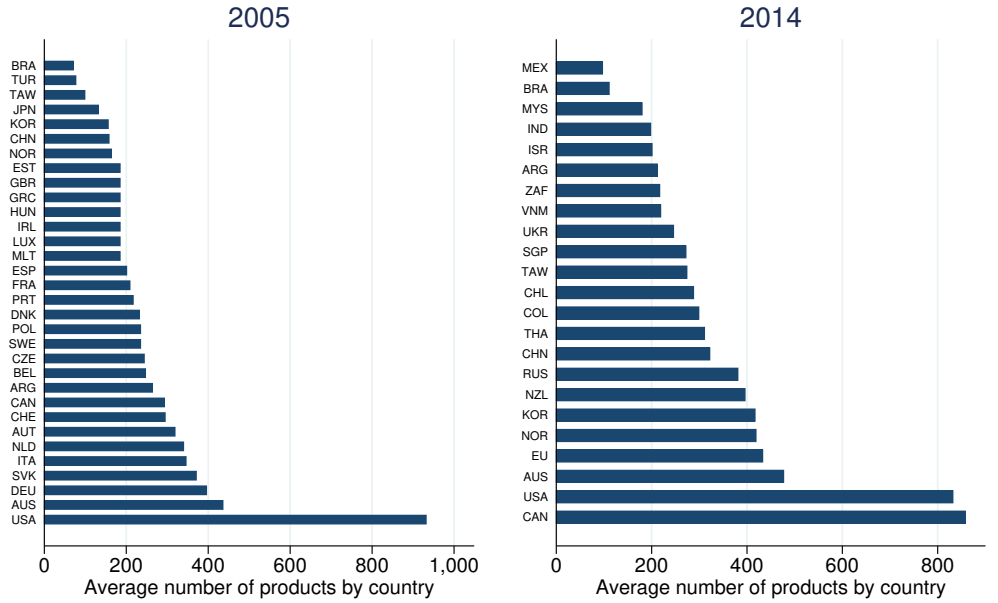
Importer	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**
ARG	506.0	506.0	506.0	506.0	506.0	506.0	506.0	506.0	506.0	506.0
AUS	457.0	457.0	457.0	457.0	457.0	457.0	457.0	457.0	457.0	457.0
AUT	492.0	505.0	548.0	624.0	767.0	767.0	767.0	767.0	767.0	767.0
BEL	456.0	479.0	473.0	537.0	732.0	732.0	732.0	732.0	732.0	732.0
BGR	201.0	234.0	308.0	465.0	612.0	612.0	612.0	612.0	612.0	612.0
BRA	335.0	338.0	444.0	351.0	361.0	364.0	363.0	371.0	217.0	236.0
CAN	322.0	322.0	322.0	322.0	322.0	322.0	322.0	322.0	322.0	322.0
CHE	426.0	438.0	480.0	547.0	707.0	707.0	707.0	707.0	707.0	707.0
CHL	154.0	166.0	305.0	163.0	173.0	186.0	181.0	226.0	234.0	241.0
CHN	170.0	180.0	340.0	214.0	244.0	259.0	255.0	281.0	382.0	412.0
COL	154.0	166.0	305.0	163.0	173.0	194.0	180.0	190.0	215.0	222.0
CYP	201.0	234.0	308.0	465.0	612.0	612.0	612.0	612.0	612.0	612.0
CZE	363.0	370.0	412.0	491.0	652.0	652.0	652.0	652.0	652.0	652.0
DEU	450.0	463.0	550.0	582.0	755.0	755.0	755.0	755.0	755.0	755.0
DNK	217.0	247.0	319.0	473.0	636.0	636.0	636.0	636.0	636.0	636.0
EGY	154.0	166.0	305.0	163.0	173.0	194.0	180.0	190.0	215.0	222.0
ESP	378.0	573.0	586.0	618.0	786.0	786.0	786.0	786.0	786.0	786.0
EST	201.0	234.0	362.0	484.0	643.0	643.0	643.0	643.0	643.0	643.0
FIN	201.0	234.0	308.0	465.0	612.0	612.0	612.0	612.0	612.0	612.0
FRA	428.0	443.0	504.0	564.0	727.0	727.0	727.0	727.0	727.0	727.0
GBR	244.0	263.0	330.0	488.0	635.0	635.0	635.0	635.0	635.0	635.0
GRC	201.0	234.0	311.0	465.0	632.0	632.0	632.0	632.0	632.0	632.0
HKG	353.0	353.0	353.0	353.0	353.0	353.0	353.0	353.0	353.0	353.0
HRV	-	-	-	-	-	-	-	-	612.0	612.0
HUN	201.0	234.0	460.0	540.0	694.0	694.0	694.0	694.0	694.0	694.0
IND	154.0	166.0	305.0	173.0	245.0	257.0	255.0	264.0	272.0	279.0
IRL	201.0	234.0	308.0	485.0	632.0	632.0	632.0	632.0	632.0	632.0
ISR	154.0	314.0	417.0	321.0	331.0	345.0	343.0	348.0	333.0	339.0
ITA	416.0	510.0	463.0	512.0	738.0	738.0	738.0	738.0	738.0	738.0
JPN	707.0	707.0	707.0	707.0	707.0	707.0	707.0	707.0	707.0	707.0
KOR	379.0	382.0	473.0	410.0	430.0	445.0	448.0	454.0	459.0	463.0
LTU	201.0	234.0	308.0	465.0	612.0	612.0	612.0	612.0	612.0	612.0
LUX	201.0	234.0	308.0	567.0	715.0	715.0	715.0	715.0	715.0	715.0
LVA	201.0	234.0	308.0	465.0	612.0	612.0	612.0	612.0	612.0	612.0
MEX	154.0	290.0	394.0	286.0	303.0	311.0	310.0	314.0	313.0	318.0
MLT	201.0	234.0	369.0	488.0	648.0	648.0	648.0	648.0	648.0	648.0
MYS	428.0	428.0	428.0	428.0	428.0	428.0	428.0	428.0	428.0	428.0
NLD	237.0	559.0	603.0	667.0	816.0	816.0	816.0	816.0	816.0	816.0
NOR	521.0	521.0	521.0	521.0	521.0	521.0	521.0	521.0	521.0	521.0
NZL	364.0	364.0	364.0	364.0	364.0	364.0	364.0	364.0	364.0	364.0
POL	343.0	359.0	451.0	525.0	696.0	696.0	696.0	696.0	696.0	696.0
PRT	326.0	357.0	396.0	485.0	655.0	655.0	655.0	655.0	655.0	655.0
ROU	201.0	234.0	308.0	465.0	612.0	612.0	612.0	612.0	612.0	612.0
RUS	154.0	166.0	320.0	389.0	413.0	425.0	422.0	428.0	450.0	471.0
SGP	154.0	166.0	305.0	163.0	199.0	216.0	214.0	224.0	235.0	242.0
SVK	229.0	260.0	419.0	501.0	654.0	654.0	654.0	654.0	654.0	654.0
SVN	201.0	234.0	308.0	465.0	612.0	612.0	612.0	612.0	612.0	612.0
SWE	254.0	289.0	337.0	485.0	636.0	636.0	636.0	636.0	636.0	636.0
THA	154.0	167.0	305.0	166.0	174.0	189.0	184.0	194.0	204.0	213.0
TUR	253.0	271.0	448.0	380.0	382.0	413.0	433.0	437.0	445.0	447.0
UKR	154.0	166.0	305.0	163.0	406.0	407.0	404.0	412.0	444.0	454.0
USA	575.0	575.0	575.0	575.0	575.0	575.0	575.0	575.0	575.0	575.0
VNM	358.0	358.0	358.0	358.0	358.0	358.0	358.0	358.0	358.0	358.0
ZAF	491.0	491.0	491.0	491.0	491.0	491.0	491.0	491.0	491.0	491.0
Total	15,631.0	17,343.0	21,198.0	22,970.0	27,929.0	28,127.0	28,094.0	28,255.0	28,952.0	29,093.0

FIGURE B.1: Average number of regulated residues per product by country of year 2005 and 2014



Source: Authors' calculations based on data on MRL regulation from Homologa, S.A.

FIGURE B.2: Average number of regulated products by country of year 2005 and 2014



Source: Authors' calculations based on data on MRL regulation from Homologa, S.A.
Note: (1) Only products with Harmonized System concordance included in this paper.
(2) From 2009, all EU member states have harmonized MRL regulation;
they thus appear as one observation in the right panel.

TABLE B.2: Many countries use Codex MRLs as default values if national regulation is missing

Country	**First default**	**Second default**
Argentina	Codex	0.01
Australia	0.01	
Brazil	Codex	
Canada	0.01	
Chile	Codex	
China	Codex	
Colombia		
Egypt	Codex	
European Union	0.01	
India	Codex	
Israel	Codex	
Japan	0.01	
Korea	Codex	
Malaysia	Codex	0.01
Mexico	Codex	
New Zealand	0.01	
Norway	0.01	
Russia	Codex	
Singapore	Codex	
South Africa	Codex	0.01
Switzerland	EU	0.01
Thailand	Codex	
Turkey	Codex	
Ukraine	Codex	
USA	0.01	
Vietnam	Codex	0.01

Note: Default MRL information were obtained from Homologa database upon request.

FIGURE B.3: Number of regulated residues

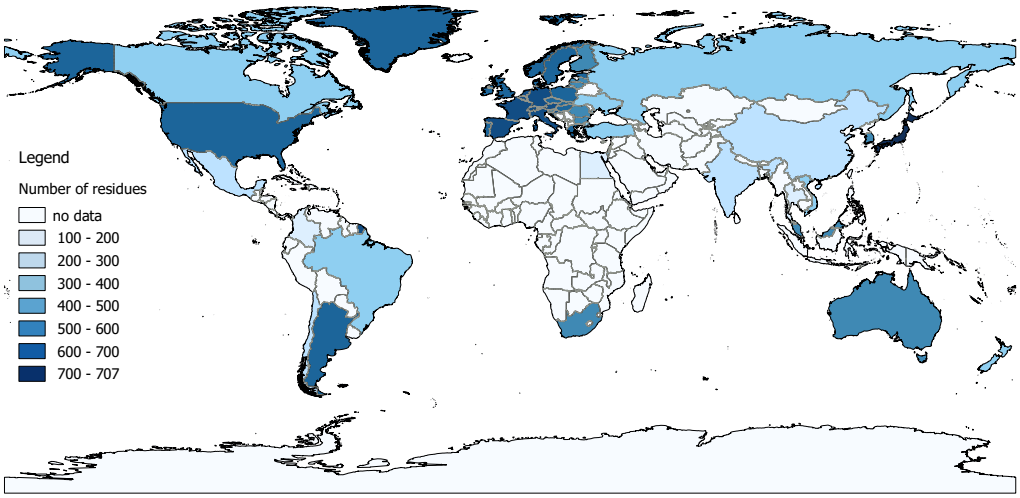
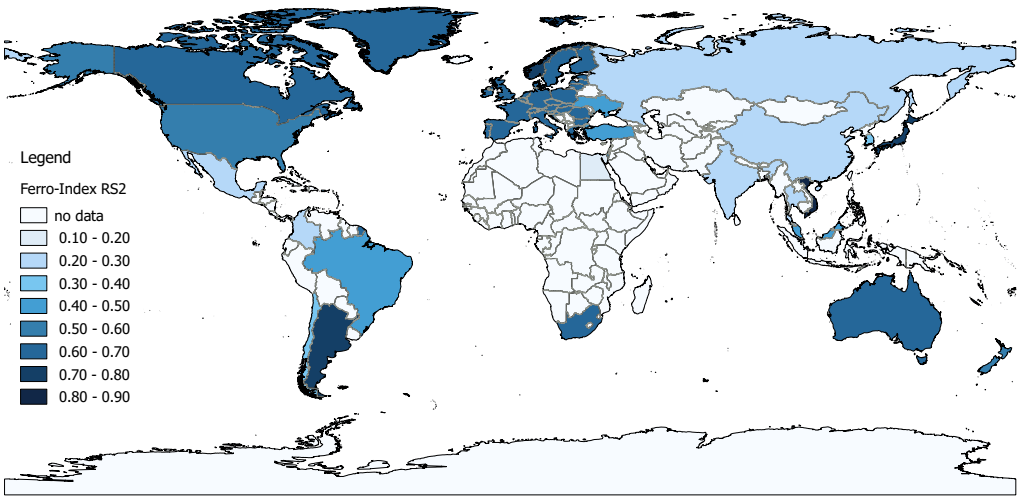


FIGURE B.4: Restrictiveness of countries (Ferro-index RS2



B.3 Coefficients of restrictiveness indices of quantile regressions across deciles

TABLE B.3: Coefficients of restrictiveness index based on Ferro *et al.* (2015) of quantile regressions across deciles.

VARIABLES	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
lnTariff	-1.023***	-0.794***	-0.440***	-0.559***	-0.126***	-0.187***	-0.0942***	-0.158***	-0.351***
	(0.101)	(0.0583)	(0.0319)	(0.0555)	(0.0143)	(0.00543)	(0.00257)	(0.00219)	(0.00375)
Ferro_RS2	-1.973***	-1.192***	-0.130***	-1.585***	-1.720***	-1.336***	-0.941***	-0.353***	0.544***
	(0.0555)	(0.0300)	(0.0160)	(0.0258)	(0.0189)	(0.00953)	(0.00416)	(0.00163)	(0.0129)
Constant	-3.690***	-2.172***	-1.654***	-0.918***	-0.365***	-0.0325***	0.0576***	0.162***	0.222***
	(0.0321)	(0.0179)	(0.0103)	(0.0152)	(0.0106)	(0.00194)	(0.00117)	(0.000935)	(0.00317)
Observations	333,258	398,922	496,251	302,624	324,140	454,470	937,776	1,369,154	1,369,960
lnTariff	-0.838***	-0.472***	-0.787***	-0.802***	-0.485***	-0.505***	-0.0647***	-0.172***	-0.471***
	(0.0937)	(0.0557)	(0.0269)	(0.0710)	(0.0414)	(0.0385)	(0.00425)	(0.00197)	(0.00455)
Ferro_RS1	-2.032***	-1.180***	0.856***	-1.740***	-1.560***	-1.284***	-1.023***	-0.227***	-1.555***
	(0.0903)	(0.0453)	(0.0190)	(0.0453)	(0.0363)	(0.0135)	(0.00688)	(0.00229)	(0.0267)
Constant	-3.348***	-2.026***	-2.215***	-0.603***	-0.281***	-0.0174**	-0.0384***	0.0619***	1.560***
	(0.0666)	(0.0338)	(0.0157)	(0.0328)	(0.0263)	(0.00842)	(0.00378)	(0.00183)	(0.0241)
Observations	343,880	413,029	571,072	284,363	297,991	345,401	675,800	1,369,871	1,369,960
lnTariff	-0.866***	-0.495***	-0.266***	-0.425***	-0.142***	-0.0540	-0.717***	-0.0753***	-0.509***
	(0.0949)	(0.0473)	(0.00896)	(0.0478)	(0.0443)	(0.0547)	(0.0198)	(0.00246)	(0.00455)
Ferro_RS0	-1.972***	-1.524***	0.0133	-2.899***	-2.779***	-2.253***	-1.243***	-0.762***	-1.036***
	(0.105)	(0.0511)	(0.0153)	(0.0484)	(0.0102)	(0.0115)	(0.0131)	(0.00305)	(0.0281)
Constant	-3.741***	-2.147***	-1.337***	-0.332***	-0.0228***	0.0706***	0.0254***	0.0809***	1.077***
	(0.0929)	(0.0453)	(0.0148)	(0.0426)	(0.00797)	(0.00843)	(0.00812)	(0.00278)	(0.0268)
Observations	365,334	445,815	906,297	251,499	256,602	286,669	358,053	1,287,270	1,369,960

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

C Appendix Chapter 4

C.1 Principal component analysis

TABLE C.1: Correlation-matrix of WGI, Source: Kaufmann *et al.* (2010)

	Corruption	Government effectiveness	Political stability	Regulatory quality	Rule of law	Voice and accountability
Corruption	1.0000					
Government effectiveness	0.950	1.0000				
Political stability	0.7873	0.7604	1.0000			
Regulatory quality	0.9066	0.9416	0.7360	1.0000		
Rule of law	0.9569	0.9603	0.7992	0.9343	1.0000	
Voice and accountability	0.8066	0.8114	0.6984	0.8369	0.8343	1.0000

TABLE C.2: Principal component rotation

Component	Eigenvalue	Difference	Proportion	Cumulative
Component 1	5	491.911	0.8757	0.8757
Component 2	.334907	0	0.0558	0.9315
Component 3	.248212	.160632	0.0414	0.9729
Component 4	.087579	.0480812	0	0.9875
Component 5	.0394978	.00370557	0.0066	1
Component 6	.0357923	-	0.0060	10.000

TABLE C.3: Principal components - eigenvectors

Variable	Component 1	Unexplained
Control of corruption	0.4214	.06687
Government effectiveness	0.4230	.06005
Political stability	0.3679	.282
Regulatory quality	0.4174	.08447
Rule of law	0.4274	.04043
Voice and accountability	0.3872	.2122

C.2 Descriptive statistics

TABLE C.4: Descriptive statistics

Variable	Source	Obs	Mean	Std. Dev.	Min	Max
Imports (current US-$, in million)	UN Comtrade	532720	8.184	67.8	0	5.12e+03
GDP_Exporter (current US-$, in million)	World Bank	52813	8.86e+5	2.24e+6	12.3	1.74e+7
GDP_Importer (current US-$, in million)	World Bank	520804	6.37e+5	1.86e+6	12.3	1.74e+7
Dist (km)	CEPII	532720	6316.499	4519.529	59.617	19812.04
Contig (=1 if common border)	CEPII	532720	.052	.222	0	1
Comlang_off (=1 if same language)	CEPII	532720	.176	.381	0	1
Colony (=1 if colonial ties)	CEPII	532720	.038	.192	0	1
lnAdvalorem_Tariff (Advalorem equivalent)	ITC	532720	.011	.048	0	.994
Control of Corruption (Index: approx. -2,5 -2,5)	World Bank	531178	.432	1.117	-1.924	2.586
Government effectiveness (Index: approx. -2,5 -2,5)	World Bank	531178	.516	1.001	-2.479	2.430
Political stability (Index: approx. -2.5 -2.5)	World Bank	531379	.103	.933	-3.324	1.938
Regulatory quality (Index: approx. -2.5 -2.5)	World Bank	531203	.482	.951	-2.675	2.247
Rule of law (Index: approx. -2.5 -2.5)	World Bank	531497	.397	1.030	-2.669	2.121
Voice and accountability (Index: approx. -2.5 -2.5)	World Bank	531491	.335	.998	-2.234	1.826

D Appendix Chapter 5

D.1 List of included products and country groups

Table D.1: Products categories

Product group	HS Code	HS Description
Bakery products	1704	Sugar confection
	1806	Chocolate & other food products containing cocoa
	1901	Malt extract, food preparations of flour etc.
	1902	Pasta, prepared or not, couscous
	1903	Tapioca and substitutes from starch in flakes, etc.
	1904	Foods prep by swell cereal, cereal n.e.s.o.i.
	1905	Bread, pastry cakes etc.
Beverages	2009	Fruit juices (& grape must), vegtables juice, no spirit
	2201	Water, natural etc., not sweetened etc., ice & snow
	2202	Water, sweetened & other non-alcoholic beverages n.e.s.o.i.
	2203	Beer made from malt
	2204	Wine of fresh grapes, grape must n.e.s.o.i.
	2205	Vermouth & other wine of fresh grapes with specific flavor
	2206	Fermented beverages n.e.s.o.i. (cider, berry, mead)
	2207	ethyl alcohol, un-denatured, n/un 80% alcohol, alcohol, denatured
	2208	ethyl alcohol, un-denatured, 80% alcohol, spirit beverages etc.
Dairy products	401	Milk and cream, not concentrated or sweetened
	402	Milk and cream, concentrated or sweetened
	403	Buttermilk, yogurt, kephir etc.
	404	Whey & milk products n.e.s.o.i.
	405	Butter and other fats and oils derived from milk
	406	Cheese and curd
Egg products	407	Birds' eggs, in the shell, fresh preserved or cooked
	408	Birds' eggs, not in shell &yolks, fresh dry, etc.
Fruits and vegetables products	2001	Vegetable, fruits, nuts, etc.
	2002	Tomatoes prepared or preserved n.e.s.o.i.
	2003	Mushrooms & truffles prepared or preserved n.e.s.o.i.
	2004	Vegetables n.e.s.o.i. prepared or preserved, frozen
	2005	Vegetables n.e.s.o.i. prepared or preserved, not frozen
	2006	Fruit/nuts/fruit-peel etc., preserved by sugar
	2007	Jams, fruit jellies, marmalade etc., cooked
	2008	Fruit, nuts etc., prepared or preserved n.e.s.o.i.
Fish products	303	Fish, frozen (no fillets)
	304	Fish fillets, other fish, fresh, chill or frozen
	305	Fish, dried, salted etc., smoked
	306	Crustaceans, live, fresh, cooked
	307	Mollusks, aquatic invertebrates n.e.s.o.i.
	1604	Prepared or preserved fish, caviar & caviar substitutes
	1605	Crustaceans and mollusks prepared or preserved
Meat products	1601	Sausages, similar prepared meat
	1602	Prepared or preserved meat, meat offal & blood n.e.s.o.i.

TABLE D.2: Importing and exporting countries

Importing countries									
Austria		Belgium		France		Germany		Italy	
Netherlands		Switzerland							
Exporting countries									
Europe				**Africa**		**America**		**Asia**	
Albania	2	Latvia	1	Ivory Coast	3	Antigua and Barbuda	2	Armenia	3
Austria	1	Lithuania	2	Egypt	3	Argentina	2	Australia	1
Belarus	2	Luxembourg	1	Ghana	3	Brazil	2	Azerbaijan	2
Belgium	1	Macedonia	2	Kenya	3	Canada	1	Bangladesh	3
Bosnia Herzegovina	2	Netherlands	1	Madagascar	3	Chile	2	China	3
Bulgaria	1	Norway	1	Mauritius	2	Colombia	2	India	3
Croatia	1	Poland	1	Morocco	3	Costa Rica	2	Indonesia	3
Cyprus	1	Romania	2	Namibia	2	Ecuador	3	Iran	2
Czech Republic	1	San Marino	1	Senegal	3	Guatemala	3	Israel	1
Denmark	1	Slovakia	1	Seychelles	2	Guyana	3	Kazakhstan	2
Estonia	1	Slovenia	1	South Africa	2	Honduras	3	South Korea	1
Faeroe Island	1	Spain	1	Tunisia	3	Mexico	2	Malaysia	2
Finland	1	Sweden	1	Uganda	3	Nicaragua	3	Pakistan	3
France	1	Switzerland	1	Peru	2			Papua New Guinea	3
Germany	1	Turkey	2	Suriname	2			Philippines	3
Greece	1	Ukraine	3	United States	1			Russia	2
Hungary	1	United Kingdom	1	Uruguay	2			Sri Lanka	3
Iceland	1							Thailand	3
Ireland	1							United Arab Emirates	1
Italy	1							Vietnam	3

Notes: numbers after country-names refer to income groups: (1) High income, (2) Upper middle income, (3) Low income.

TABLE D.3: Estimation results by product using ppml and IFS-certification in neighbouring countries as instruments using GMM and the Control function approach (cfa).

	Bakery			**Beverages**			**Egg**		**Meat**		
VARIABLES	ppml	IV (cfa)	IV(gmm)	ppml	IV (cfa)	IV(gmm)	ppml	IV (cfa)	ppml	IV (cfa)	IV(gmm)
IFS_lag	0.885***			0.415***			0.316**		0.543***		
	(0.021)			(0.026)			(0.134)		(0.027)		
lnGDP_Importer	0.703***	0.727***	0.752***	0.799***	0.590***	0.856***	1.165***	0.737***	0.723***	0.821***	0.744***
	(0.045)	(0.090)	(0.106)	(0.062)	(0.101)	(0.143)	(0.211)	(0.054)	(0.062)	(0.130)	(0.141)
lnGDP_Exporter	0.291***	0.443***	0.200***	0.590***	0.924***	0.511***	0.553***	0.343***	0.413***	0.906***	0.554***
	(0.022)	(0.059)	(0.074)	(0.024)	(0.077)	(0.066)	(0.109)	(0.040)	(0.029)	(0.097)	(0.097)
lnDist	-0.440***	-0.841***	-0.438**	-1.334***	-1.023***	-1.352***	-2.072***	-0.845***	-0.913***	-1.463***	-0.917***
	(0.087)	(0.192)	(0.195)	(0.112)	(0.205)	(0.240)	(0.281)	(0.127)	(0.109)	(0.303)	(0.243)
lnTariff	-0.758	3.038***	2.339***	2.587***	5.285***	2.852***	-0.407	0.005	-0.194	0.605	0.261
	(0.696)	(0.713)	(0.782)	(0.513)	(0.460)	(0.629)	(0.782)	(0.158)	(0.547)	(0.488)	(0.271)
Comlang_ethno	0.929***	0.425	1.010***	0.876***	-0.234	0.970***	0.273	0.760***	1.311***	0.378	1.368***
	(0.126)	(0.285)	(0.279)	(0.153)	(0.377)	(0.333)	(0.249)	(0.159)	(0.163)	(0.451)	(0.374)
Colony	-0.287	0.461	-0.275	-0.173	0.424	-0.196	0.851**	0.206	-0.099	1.496**	0.011
	(0.195)	(0.346)	(0.446)	(0.189)	(0.404)	(0.479)	(0.387)	(0.221)	(0.167)	(0.712)	(0.397)
Contig	0.581***	0.646**	0.530**	-0.088	0.942***	-0.172	0.653***	0.241	0.358**	0.647	0.410
	(0.114)	(0.251)	(0.246)	(0.182)	(0.300)	(0.362)	(0.168)	(0.152)	(0.147)	(0.439)	(0.327)
RTA	1.620***	1.136***	1.393***	0.980***	0.641**	0.804	-0.317	1.470***	2.011***	0.366	2.391***
	(0.168)	(0.263)	(0.342)	(0.283)	(0.327)	(0.563)	(0.496)	(0.267)	(0.249)	(0.580)	(0.542)
IFS (IV)		0.832***	1.057***		0.487***	0.504***		0.622***		0.406***	0.445***
		(0.053)	(0.078)		(0.070)	(0.073)		(0.047)		(0.107)	(0.107)
Constant	-26.230***	-30.763***	-25.966***	-34.828***	-38.192***	-34.557***	-44.413***	-27.388***	-28.950***	-44.605***	-33.042***
	(1.336)	(2.912)	(3.265)	(1.919)	(3.241)	(4.423)	(5.982)	(1.827)	(2.076)	(4.430)	(4.805)
Observations	9,996	9,556	9,556	14,049	12,915	12,915	1,008	56,043	2,548	2,534	2,534
R-squared	0.309			0.167			0.488		0.443		

	Fish			**F&V**			**Dairy**	
VARIABLES	ppml	IV (cfa)	IV(gmm)	ppml	IV (cfa)	IV(gmm)	ppml	IV(cfa)
IFS_lag	0.532***			0.567***			0.551***	
	(0.030)			(0.019)			(0.028)	
lnGDP_Importer	0.735***	0.930***	0.705***	0.837***	0.977***	0.864***	0.658***	1.068***
	(0.036)	(0.115)	(0.079)	(0.046)	(0.098)	(0.106)	(0.053)	(0.144)
lnGDP_Exporter	0.221***	0.825***	0.389***	0.315***	0.540***	0.058	0.489***	-0.270*
	(0.026)	(0.094)	(0.064)	(0.025)	(0.066)	(0.185)	(0.038)	(0.160)
lnDist	-0.681***	-0.736***	-0.546**	-0.874***	-0.619***	-0.928***	-0.685***	-0.488**
	(0.101)	(0.249)	(0.216)	(0.108)	(0.205)	(0.251)	(0.117)	(0.223)
lnTariff	1.419**	3.630**	1.689*	-0.366	-1.609**	1.033*	-3.336***	-2.355***
	(0.561)	(1.516)	(0.869)	(0.232)	(0.741)	(0.565)	(0.547)	(0.425)
Comlang_ethno	0.568***	-0.798	0.523	0.800***	0.890**	0.791***	0.533***	0.650*
	(0.148)	(0.505)	(0.354)	(0.117)	(0.394)	(0.260)	(0.134)	(0.367)
Colony	0.807***	0.781	0.988**	-0.141	0.366	0.011	0.315	0.199
	(0.185)	(0.492)	(0.442)	(0.166)	(0.514)	(0.419)	(0.213)	(0.574)
Contig	0.136	1.235***	0.345	0.276*	0.438*	0.161	0.352**	0.737*
	(0.121)	(0.463)	(0.283)	(0.148)	(0.257)	(0.300)	(0.151)	(0.386)
RTA	0.332**	-0.206	0.646*	2.226***	1.142***	1.658***	2.427***	1.734***
	(0.156)	(0.313)	(0.354)	(0.232)	(0.332)	(0.393)	(0.310)	(0.470)
IFS (IV)		-0.326*	0.220***		0.593***	0.862***		1.566***
		(0.194)	(0.084)		(0.093)	(0.210)		(0.156)
Constant	-23.497***	-43.222***	-26.332***	-30.206***	-40.267***	-25.269***	-28.305***	-22.566***
	(1.369)	(3.802)	(3.183)	(1.440)	(3.294)	(5.096)	(1.706)	(5.481)
Observations	13,279	12,740	12,740	12,992	11,760	11,760	5,586	5,460
R-squared	0.164			0.225			0.248	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Because estimations were based on the sectoral level, the method of Baier-Bergtrand was used to account for multilateral resistance. The command `ivpoisson` was used in Stata. CFA refers to control function approach, gmm to general methods of moments. First stage results are omitted here.

E Appendix Chapter 6

E.1 The Synthetic Control Method (SCM)

E.1.1 The synthetic control

The following exposition is based on Abadie *et al.* (2003), Abadie *et al.* (2010) and Abadie *et al.* (2015). For countries $i = 1, \ldots, J+1$ and periods $t = 1, \ldots, T$, suppose country i is exposed to the intervention (Hartz reforms) at $T_0 \in (1, T)$. The observed outcome for country i at time t is,

$$Y_a = Y_a^N + \alpha S_a \tag{E.1}$$

where Y_a^N is the outcome for country i at time j in the absence of the intervention, the binary indicator variable S_{it} denotes the existence of the intervention (Hartz reforms) taking the value 1 if $i = 1$ at $t > T_0$ and $\alpha_t t$ is the effect of the intervention for country i at time t. Thus, country i is exposed to the intervention in periods $T_0 + 1$ to T. We assume that the passage of the Hartz reforms had no effect on the outcome in Germany before the implementation period. We restrict the donor pool to countries that were not exposed to the Hartz reforms over the period $t = 1, \ldots, T$, and assume that the outcomes of the untreated countries were not affected by the passage of the Hartz reforms in Germany.

Indexing the exposed country Germany as country 1, we want to estimate $(\alpha_{1T_0+1}, \ldots, \alpha_{1T})$. From Equation E.1 we note that $\alpha_{1t} = Y_{1t} - Y_{1t}^N$ for $t \in \{T_0 + 1, \ldots, T\}$, and while Y_{1t} is observed Y_{1t}^N is unobserved. We, therefore, need to estimate Y_{1t}^N.

Suppose Y_{it}^N is given by the model

$$Y_{it}^N = \delta_t + \theta_t \mathbb{Z}_t + \lambda_t \mu_i + \varepsilon_{it} \tag{E.2}$$

where, δ_t is an unknown common factor constant across countries, $\mathbb{Z}_t$ is a $(r \times 1)$ vector of observed covariates (not affected by the intervention), θ_t is a $(1 \times r)$ vector of unknown parameters, λ_t is a $(1 \times F)$ vector of unobserved time-varying common factors, μ_i is a $(F \times 1)$ vector of unknown unit specific factors, and ε_{it} are the unobserved transitory shocks at the country level with zero mean.

Consider a $(J \times 1)$ vector of weights $\mathbb{W} = (w_2, \ldots, w_{J+1})'$ such that $\{w_j \geq 0 | j = 2, \ldots, J+1\}$ and $\sum_{j=2}^{J+1} w_j = 1$. Each value of the vector $\mathbb{W}$ represents a weighted average of the control countries and, hence, a potential synthetic control. Abadie, Diamond, and Hainmueller (2010) show that, there exist $\mathbb{W}^* = (w_2^*, \ldots, w_{J+1}^*)'$ such that, $Y_{1t}^N = \sum_{j=2}^{J+1} w_j^* Y_{jt}$, $t = 1, \ldots, T_0$ and $\mathbb{Z}_1 = \sum_{j=2}^{J+1} w_j^* \mathbb{Z}_j$ (that is, pre-intervention

matching with respect to the outcome variable as well as the covariates, henceforth referred to as predictors), then under standard conditions we can use,

$$\hat{\alpha}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt} \qquad t \in \{T_0 + 1, \ldots, T\} \tag{E.3}$$

as an estimator for α_{1t}. The term $\sum_{j=2}^{J+1} w_j^* Y_{jt}$ on the right-hand-side of Equatin E.3 is simply the weighted average of the observed outcome of the control countries for $t \in \{T_0 + 1, \ldots, T\}$ with weights $\mathbb{W}^*$. The procedure to obtain $\mathbb{W}^*$ is discussed in Abadie, Diamond, and Hainmueller (2010).

It is important to note, as Abadie *et al.* (2010) argue, equation (E.2) is a generalisation and that the traditional regression-based difference-in-difference model can be obtained if we impose that λ_t be constant for all t. Thus, unlike the traditional regression-based difference-in-difference model that restricts the effects of the unobservable confounders to be time-invariant so that they can be eliminated by taking time differences, this model allows the effects of such unobservables to vary with time. In particular, Abadie *et al.* (2010) show that a synthetic control can fit $\mathbb{Z}_1$ and a long set of pre-intervention outcomes, $Y_{11}, \ldots, Y_{1T_0}$, only as long as it fits $\mathbb{Z}_1$ and μ_1 (unknown factors of the exposed unit).

E.1.2 Inference

Once an optimal weighting vector $\mathbb{W}^*$ is chosen, the "synthetic Germany" is obtained by calculating the weighted average of the donor pool. The post-intervention values of the synthetic serve as our counterfactual outcome for Germany. Following Abadie *et al.* (2010) we calculate the ratio of post-intervention to pre-intervention Root Mean Square Prediction Error or RMSPE (the square root of the squared difference between the actual outcome and the synthetic outcome), denoted by Δ_{DUE}. This ratio puts the magnitude of post intervention gap (between the actual and the synthetic outcome) in the context of the pre-intervention fit (between the actual and the synthetic outcome): the larger the ratio the greater is the impact of the intervention.

To formally test the significance of this estimate, we apply the permutations test suggested by Bertrand *et al.* (2004), Buchmueller *et al.* (2011), Abadie *et al.* (2010) and Bohn *et al.* (2014). First, for each country in the donour pool, we carry out an SCM estimate as if the country had passed the Hartz reforms in 2003 (i.e., apply a fictitious policy intervention). We can then calculate the post-pre RMSPE ratio for each of these countries. The distribution of these "placebo" post-pre RMSPE ratios, $F(\Delta)$, then provides the equivalent of a sampling distribution for Δ_{DUE} that is used to calculate the p-value (Bohn *et al.* (2014), Munasib *et al.* (2015)). Note that this answers the question, how often would we obtain an effect of the Hartz reforms of a magnitude as large as that of Germany if we had chosen a country at random, which is the fundamental question of inference (Bertrand *et al.* (2004); Buchmueller *et al.* (2011); Abadie *et al.* (2010).

Following Abadie *et al.* (2010) we provide two additional tests of inference. First we examine the ranking of the magnitude of the post-pre RMSPE ratio of Germany vis-à-vis those of the placebos. If Germany is ranked among the top, then we consider it significant, the rationale being that for the treatment effect to be significant the placebo effects, in general, should not be larger than the actual effect estimated

for the exposed unit. And, finally, Abadie *et al.* (2010) produce a statistic that is obtained by dividing the rank of the post-pre RMSPE ratio by one plus the size of the donour pool; this is the probability of obtaining a post-pre RMSPE ratio as large as the treated if one were to assign the intervention at random in the data. We call this statistic "donour probability" and report it for each estimate.

E.2 Descriptive statistics

Table E.1: List of Countries

All OECD (35 countries)	Donour pool (26 countries)
Country	Country
Australia	Australia
Austria	Belgium
Belgium	Canada
Canada	Switzerland
Switzerland	Denmark
Chile	Spain
Czech Republic	Estonia
Germany	Finland
Denmark	France
Spain	United Kingdom
Estonia	Greece
Finland	Ireland
France	Iceland
United Kingdom	Israel
Greece	Italy
Hungary	Japan
Ireland	Korea
Iceland	Luxembourg
Israel	Mexico
Italy	Netherlands
Japan	Norway
Korea	New Zealand
Luxembourg	Portugal
Latvia	Sweden
Mexico	Turkey
Netherlands	United States
Norway	
New Zealand	
Poland	
Portugal	
Slovak Republic	
Slovenia	
Sweden	
Turkey	
United States	

OECD country list reported as of 2016. Chile, Estonia, Israel and Slovenia joined in 2010; Latvia joined in 2016.

Table E.2: Summary Statistics 1991-2015

	All OECD countries				Donour pool		Germany	
	Mean	Std	Min	Max	Mean	Std	Mean	Std
Outcome variables (1991-2015)								
Unemployment rate	7.78	4.09	1.47	27.47	7.43	4.08	7.8	1.79
Unemployment rate – men	7.43	3.97	1.14	25.6	7.07	3.87	7.59	1.83
Unemployment rate – women	8.32	4.75	1.53	31.62	7.99	4.94	8.1	1.98
Employment in working age population (%)	65.54	7.86	44.26	85.76	66.46	8.2	67.76	3.52
Employment in working age population - men (%)	65.17	6.85	45.94	85.55	66.34	6.98	62.07	2.39
Employment in working age population - women (%)	48.08	10.29	20.69	76.8	49.06	11.28	46.43	3.37
Labour force participation rate: age 15-64	70.98	7.09	49.64	87.96	71.7	7.47	73.46	2.85
Labour force participation rate: men	70.31	5.63	58.89	87.78	71.29	5.82	67.16	1.7
Labour force participation rate: women	52.26	9.99	23.3	79.28	53.09	10.98	50.49	2.76
Labour force participation rate: age 15-24	49.41	13.49	24.28	80.06	51.84	13.55	52.1	3.12
Labour force participation rate: age 25-54	82.75	6.45	57.82	92.49	82.24	6.75	85.73	1.73
Labour force participation rate: age 55-64	50.98	15.01	17.53	88.72	53.84	14.58	51.25	10.57
Predictors (1991-2002)								
Population weighted distance (km)	7723.34	4656	110.88	18525.52	7870.92	4637.17	5428.44	4088.8
Education expenditure (% of GNI)	5.03	1.36	2.14	8.28	5.12	1.46	4.3	0.04
Domestic credit provided by banking sector (% of GDP)	93.3	52.56	11.21	305.01	101.98	54	130.24	15.7
Openness at 2005 constant prices (%)	68.46	40.17	16.65	265.25	66.38	44.13	51.18	8.88
Consumption share of PPP converted PC GDP ('05 USD)	67.63	7.04	46.39	79.21	66.79	7.42	70.03	0.84
Government share of PPP converted PC GDP ('05 USD)	7.23	2.29	3.46	17	6.95	1.94	6.1	0.26
Investment share of PPP converted PC GDP ('05 USD)	23.27	5.33	9.98	45.76	23.76	5.5	23.02	1.13
Per capita GDP (WDI, constant 2005 USD)	17435.82	10649.48	2291.46	48473.6	20519.83	9915.98	21633.31	1071.12
Share of population age 65 and over	13.04	3.29	4.33	18.66	13.14	3.45	15.77	0.63
Years of schooling	9.89	1.51	5.1	12.73	9.77	1.58	9.64	0.58
Population density (per sq mile)	121.73	113.22	2.24	476.83	127.17	126.75	228.49	2.08
Urban population (% of total)	74.1	11.14	48.55	97.19	76.64	10.47	73.18	0.06
% of urban population in the largest city	26.73	14.24	5.62	67.09	27.24	14.25	5.74	0.1
Infant mortality rate (per 1,000 live births)	8.47	7.38	2.7	56.5	8.08	8.18	5.14	0.79
Domestic credit to private sector (% of GDP)	76.25	44.07	7.17	227.77	84.45	44.57	106.91	11.1
Trade union density (%)	32.75	20.43	7.9	92.5	34.71	22.42	24.27	0.83

The data period is 1991-2015.
Intervention (beginning of the Hartz reforms) is in 2003. Thus, the pre-intervention period is 1991-2002.
No outcome variables is available for all 35 countries for the entire study period. The same is also true from a few of the predictors.
In the SCM analysis reported in the rest of the tables, logarithms of population weighted distance, per capita GDP and population densities were used.

E.3 SCM Estimate of the Impact of Hartz Reforms on Labour Force Participation Rate in Germany

FIGURE E.1: SCM Estimate of the Impact of Hartz Reforms on labour Force Participation Rate in Germany (corresponds to Table E.3)

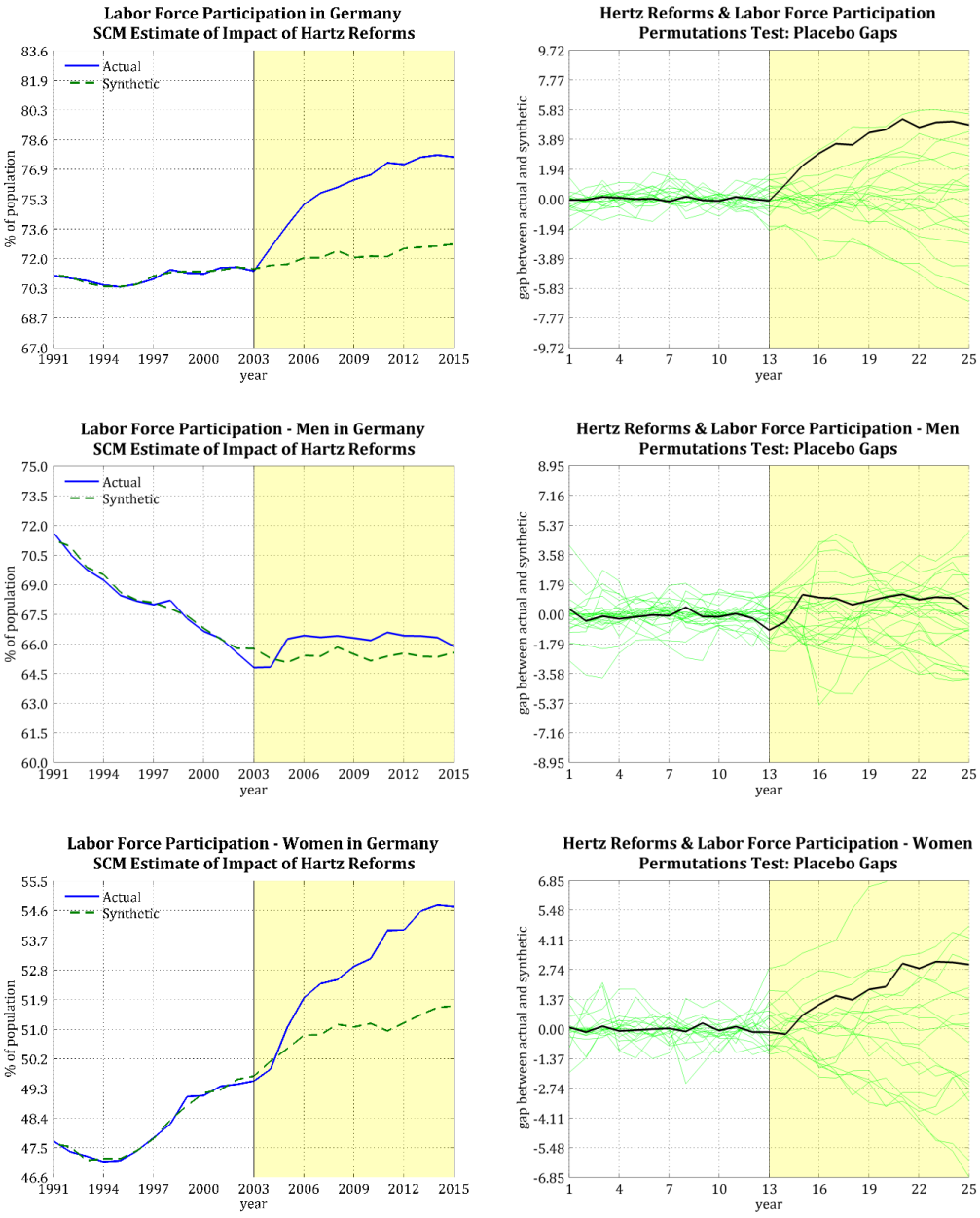


TABLE E.3: SCM Estimates of the Impact of the Hartz Reforms on labour Force Participation Rate in Germany

	labour force participation rate				labour force participation rate			
	All	Men	Women		All	Men	Women	Actual
Panel A: Estimation statistics				*Panel C: Pre-int characteristic match*	Synthetic Germany			Germany
APEMR	0	0	0	Population weighted distance	8.79	8.54	8.74	8.74
Pre-intervention RMSPE	0.1	0.24	0.12	Education expenditure	5.16	4.33	4.64	4.3
P-value: Post-pre RMSPE ratio	0	0.37	0	Domestic credit by banking sector	114.25	70.44	141.22	130.24
post/pre RMSPE ratio rank	1	11	1	Openness at 2005 constant prices	51.18	51.51	53.18	51.18
Donour probability	0.04	0.41	0.04	Consumption share of GDP	70.03	70.81	69.8	70.03
Estimated impact	3.91	0.79	1.92	Government share of GDP	6.13	6.47	6.43	6.1
Panel B: w-weights				Investment share of GDP	23.02	21.37	24.23	23.02
Belgium	0	0	0.1	Per capita GDP	9.72	9.16	9.89	9.98
Canada	0.12	0	0	Population 65 and older	12.27	13.27	15.64	15.77
Switzerland	0.15	0	0	Years of schooling	9.59	8.22	9.28	9.64
Spain (España)	0	0.05	0					
Estonia	0	0.17	0.03					
France	0	0.18	0					
Great Britain	0	0	0.33					
Ireland	0	0	0.04					
Israel	0.04	0	0					
Italy	0.1	0.35	0.16					
Japan	0	0	0.19					
Mexico	0.18	0	0					
Portugal	0.08	0	0.16					
Sweden	0.1	0	0					
Turkey	0.04	0.25	0					
United States	0.18	0	0					

The data period is 1991-2015.
Intervention (beginning of the Hartz reforms) is in 2003. Thus, the pre-intervention period is 1991-2002.
APEMR = absolute prediction error to mean ratio, Estimated impact = annual average post-intervention gap between actual and synthetic.
Population weighted distance and per capita GDP are in logarithms.
Only countries with weights exceeding 0.01 were reported. The countries with weights smaller than or equal to 0.01 are: Australia, Denmark, Finland, Greece, Iceland, South Korea, Luxembourg, Netherlands, Norway and New Zealand.

E.4 SCM Estimate of the Impact of Hartz Reforms on % of Working Age Population that is Employed in Germany

FIGURE E.2: SCM Estimate of the Impact of Hartz Reforms on % of Working Age Population that is Employed in Germany (corresponds to Table E.4)

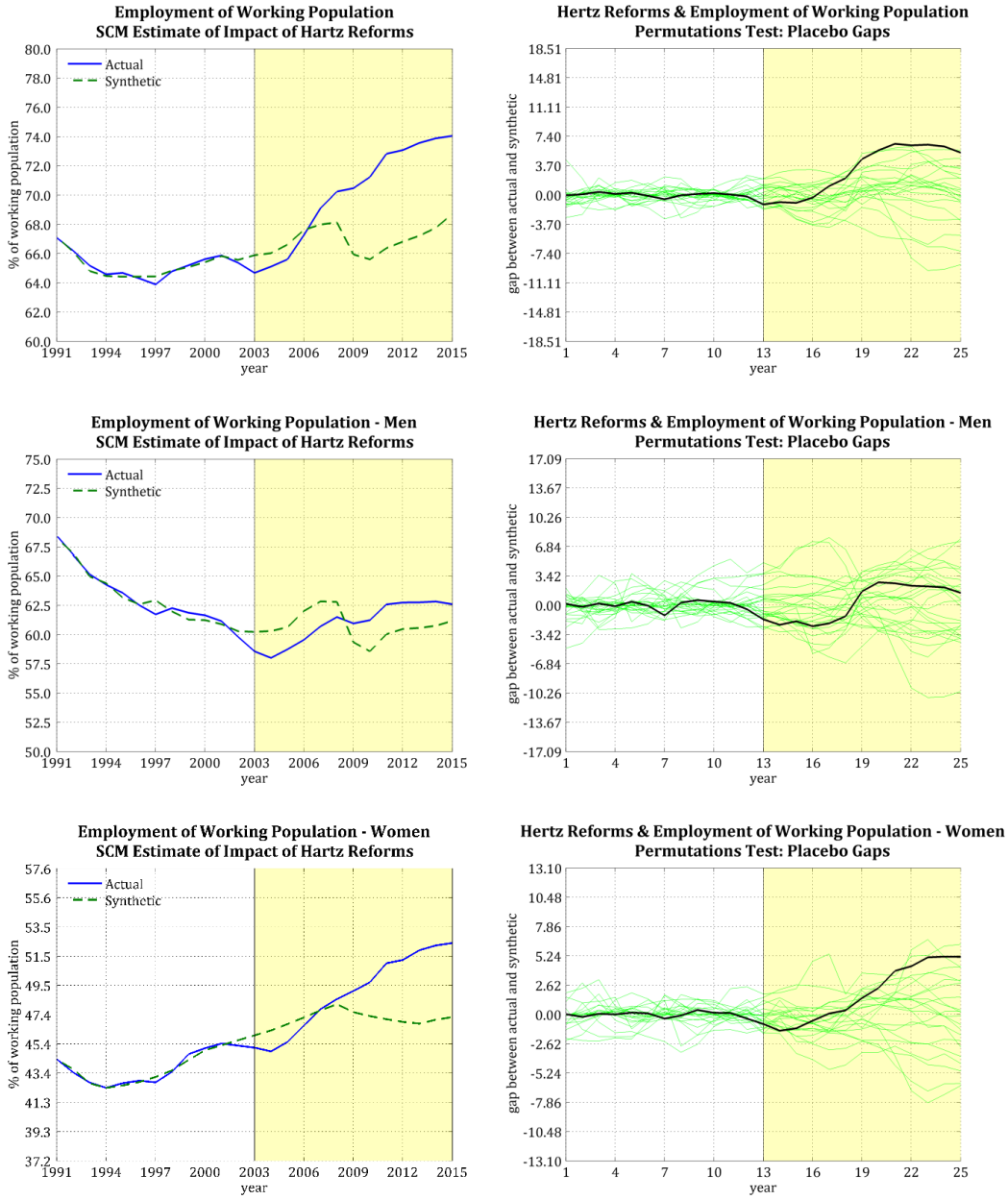


TABLE E.4: SCM Estimates of the Impact of the Hartz Reforms on % of Working Age Population Employed in Germany

	% of working age population employed				% of working age population employed			
Panel A: Estimation statistics	All	Men	Women	*Panel C: Pre-int characteristic match*	All Synthetic Germany	Men	Women	Actual Germany
APEMR	0	0.01	0	Population weighted distance	8.79	8.9	8.57	8.74
Pre-intervention RMSPE	0.23	0.47	0.22	Education expenditure	4.76	4.51	4.77	4.3
P-value: Post-pre RMSPE ratio	0	0.22	0	Domestic credit by banking sector	130.24	130.23	130.24	130.24
post/pre RMSPE ratio rank	1	7	1	Openness at 2005 constant prices	51.18	101.66	51.18	51.18
Donour probability	0.04	0.26	0.04	Consumption share of GDP	68.19	66.38	68.38	70.03
Estimated impact	3.46	0.39	2.05	Government share of GDP	6.65	6.69	7.36	6.1
Panel B: w-weights				Investment share of GDP	23.96	24.92	24.05	23.02
Belgium	0	0.51	0.1	Per capita GDP	9.86	9.54	9.96	9.98
Estonia	0.1	0.3	0	Population 65 and older	14.79	15.4	15.53	15.77
Finland	0	0	0.2	Years of schooling	9.61	10.32	9.31	9.64
France	0	0	0.19					
Great Britain	0	0	0.12					
Greece	0	0	0.09					
Iceland	0.13	0	0					
Italy	0.3	0	0.11					
Japan	0.24	0.19	0.18					
Luxembourg	0.02	0	0					
Sweden	0.14	0	0					
Turkey	0.05	0	0					
United States	0.02	0	0					

The data period is 1991-2015.

Intervention (beginning of the Hartz reforms) is in 2003. Thus, the pre-intervention period is 1991-2002.

APEMR = absolute prediction error to mean ratio, Estimated impact = annual average post-intervention gap between actual and synthetic.

Population weighted distance and per capita GDP are in logarithms.

Only countries with weights exceeding 0.01 were reported. The countries with weights smaller than or equal to 0.01 are: Australia, Canada, Denmark, Ireland, Israel, Mexico, Netherlands, New Zealand, Norway, Portugal, South Korea, Spain (España) and Switzerland.

E.5 SCM Estimate of the Impact of Hartz Reforms on labour Force Participation Rate by Age Groups in Germany

FIGURE E.3: SCM Estimate of the Impact of Hartz Reforms on Labour Force Participation Rate by Age Groups in Germany (corresponds to Table E.3)

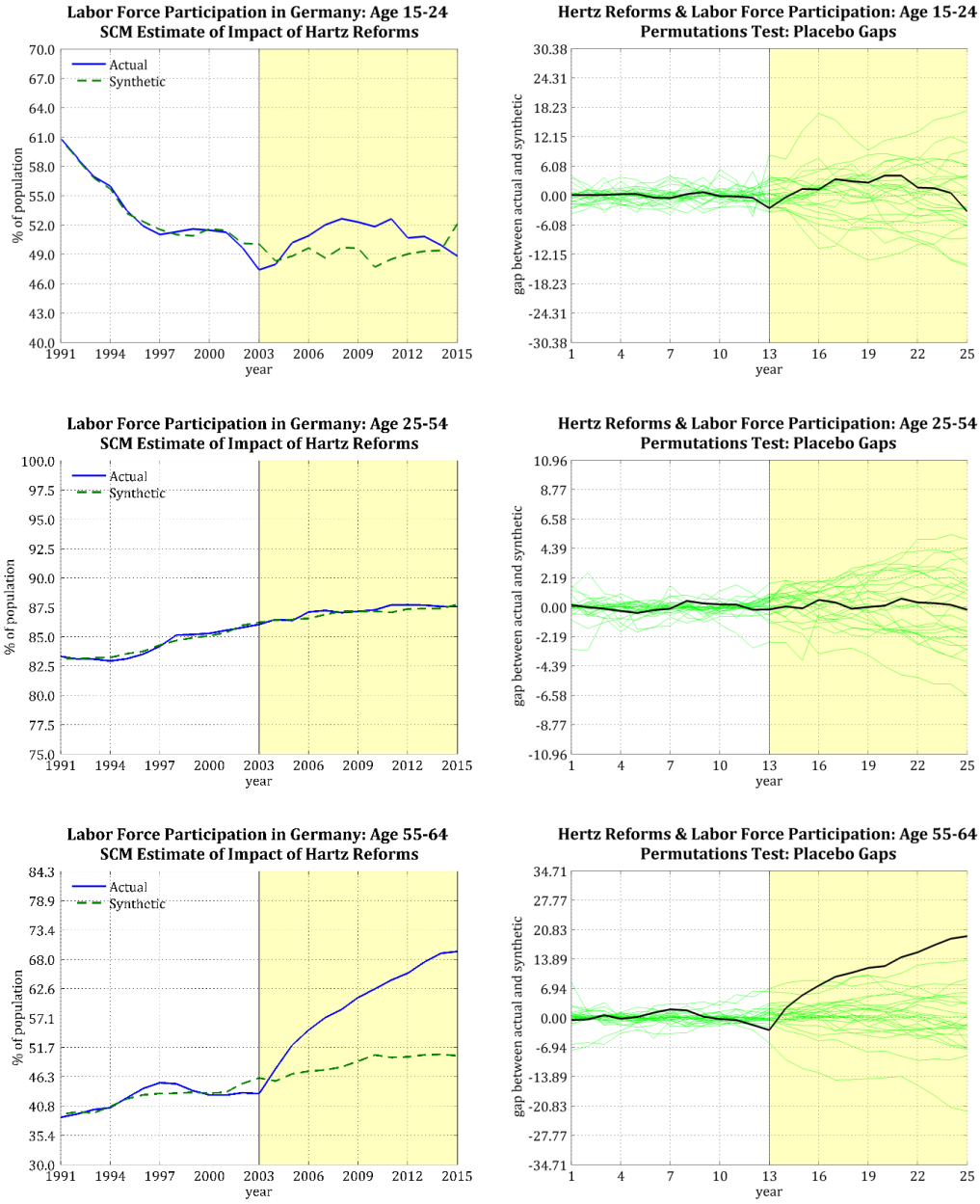


TABLE E.5: SCM Estimates of the Impact of the Hartz Reforms on Age Distribution of labour Force Participation Rate in Germany

	labour force participation rate by age groups				labour force participation rate by age groups			
	15-24	25-54	55-64		15-24	25-54	55-64	Actual
Panel A: Estimation statistics				*Panel C: Pre-int characteristic match*	Synthetic Germany			Germany
APEMR	0.01	0	0.02	Population weighted distance	8.69	8.7	8.7	8.74
Pre-intervention RMSPE	0.36	0.26	1.03	Education expenditure	4.98	5.36	3.99	4.3
P-value: Post-pre RMSPE ratio	0.11	0.89	0.04	Domestic credit by banking sector	100.6	130.24	83.22	130.24
post/pre RMSPE ratio rank	4	25	2	Openness at 2005 constant prices	113.86	62.63	68.78	51.18
Donour probability	0.15	0.93	0.07	Consumption share of GDP	63.88	67.88	71.38	70.03
Estimated impact	1.65	0.18	12.03	Government share of GDP	5.94	6.47	7.54	6.1
Panel B: w-weights				Investment share of GDP	22.4	23.07	21.4	23.02
Belgium	0	0	0.03	Per capita GDP	9.92	9.97	9.3	9.98
Canada	0	0.46	0	Population 65 and older	13.97	13.45	12.6	15.77
Switzerland	0.26	0.25	0	Years of schooling	9.73	10.01	9.14	9.64
Estonia	0.15	0	0					
Greece	0	0.1	0.41					
Iceland	0.08	0.06	0					
Luxembourg	0.26	0	0.12					
Netherlands	0	0.02	0					
Norway	0	0	0					
New Zealand	0	0	0.27					
Portugal	0	0.08	0					
Sweden	0.16	0	0					
Turkey	0.07	0	0.18					
United States	0	0.03	0					

The data period is 1991-2015.
Intervention (beginning of the Hartz reforms) is in 2003. Thus, the pre-intervention period is 1991-2002.
APEMR = absolute prediction error to mean ratio, Estimated impact = annual average post-intervention gap between actual and synthetic.
Population weighted distance and per capita GDP are in logarithms.
Only countries with weights exceeding 0.01 were reported. The countries with weights smaller than or equal to 0.01 are: Australia, Denmark, Finland, France, Great Britain, Ireland, Israel, Italy, Japan, Mexico, Norway, South Korea and Spain (España).

TABLE E.6: Summary of SCM Estimates of the Impact of the Hartz Reforms on labour Market Outcomes in Germany - Set of Alternative Predictors

	% of working age population employed			labour force participation rate			labour force participation rate by age groups			
	All	Men	Women	All	Men	Women	15-24	25-54	55-64	
APEMR	0	0	0	0	0	0	0.01	0	0.02	
Pre-intervention RMSPE	0.3	0.38	0.21	0.09	0.27	0.11	0.41	0.24	1.03	
P-value: Post-pre RMSPE ratio	0	0.26	0.04	0	0.7	0	0.07	0.7	0.04	
post/pre RMSPE ratio rank	1	8	2	1	20	1	3	20	2	
Donour probability	0.04	0.3	0.07	0.04	0.74	0.04	0.11	0.74	0.07	
Estimated impact	3.27	0.55	2.42	3.09	0.13	1.9	3.39	-0.07	12.65	
Top five countries with the largest weights	Japan France Estonia Portugal Italy	France Italy Japan Finland USA	France Italy Japan Finland USA	CHE Israel Italy Turkey Portugal	France Estonia Italy Japan LUX	Japan Portugal Belgium Italy Ireland	LUX CHE Sweden Estonia USA	Norway CHE NLD Portugal Denmark	Greece NZL Turkey France LUX	
Pre-int characteristic match	Synthetic Germany									Actual Germany
Population density	4.62	4.57	4.62	5.09	4.47	5	4.21	4.61	4.01	5.43
Urban population	73.2	73.34	73.22	73.17	73.19	73.18	77.16	71.31	69.73	73.18
Urban population in largest city	31.89	32.01	23.38	27.84	28.34	31.89	23.16	23.54	34.36	5.74
Infant mortality rate	5.88	11.69	5.27	8.87	7.18	5.76	6.15	5.53	14.27	5.14
Education expenditure	4.8	5	4.86	4.73	5.1	4.59	5.16	5.62	4.05	4.3
Domestic credit to private sector	108.88	58.44	106.9	107.15	79.22	106.91	106.9	95.05	56.6	106.91
Per capita GDP	9.77	9.15	10	9.83	9.4	9.82	10.06	10.06	9.23	9.98
Population 65 and older	15.21	13.36	15.15	13.6	15.09	14.67	14.45	15.03	12.27	15.77
Years of schooling	9.49	9.08	9.34	9.33	9.66	9.52	10.44	9.67	8.91	9.64
Union density	24.28	24.28	28.96	26.3	14.63	32.09	36.11	36.1	24.97	24.28

The data period is 1991-2015.
Intervention (beginning of the Hartz reforms) is in 2003. Thus, the pre-intervention period is 1991-2002.
APEMR = absolute prediction error to mean ratio, Estimated impact = annual average post-intervention gap between actual and synthetic.
The donor pool is the same as the main estimates in Tables 2-4.
For brevity, the following ISO codes were used in the table: CHE for Switzerland, LUX for Luxembourg, NLD for the Netherlands and, of course, USA for the United States.

TABLE E.7: Summary of SCM Estimates of the Impact of the Hartz Reforms on labour Market Outcomes in Germany - Alternative Donour Pools

	% of working age population employed			labour force participation rate			labour force participation rate by age groups		
	All	Men	Women	All	Men	Women	15-24	25-54	55-64
APEMR	0	0.01	0	0	0.01	0	0.01	0	0.02
Pre-intervention RMSPE	0.28	0.46	0.2	0.1	0.55	0.21	0.46	0.33	1.14
P-value: Post-pre RMSPE ratio	0	0.31	0	0	0.81	0.04	0.08	0.88	0.04
post/pre RMSPE ratio rank	1	9	1	1	22	2	3	24	2
Donour probability	0.04	0.35	0.04	0.04	0.85	0.08	0.12	0.92	0.08
Estimated impact	3.41	0.78	3.49	3.31	0.26	1.9	3.36	-0.13	13.43
Excluded from donor pool	Italy	Belgium	Finland	USA	Italy	G. Britain	CHE	Canada	Greece
Top five countries with the largest weights	France Japan Estonia Sweden Iceland	France Estonia Turkey Italy Japan	Italy Portugal USA Sweden Belgium	CHE Japan Mexico Portugal Canada	France Japan Estonia Belgium LUX	Italy Belgium USA Canada France	LUX G. Britain Iceland USA Estonia	CHE Norway Greece Portugal Denmark	Turkey Belgium N Zealand S Korea Japan

The data period is 1991-2015.
Intervention (beginning of the Hartz reforms) is in 2003. Thus, the pre-intervention period is 1991-2002.
APEMR = absolute prediction error to mean ratio, Estimated impact = annual average post-intervention gap between actual and synthetic.
The donor pool is the same as the main estimates in Tables 2-4.
For each outcome the donor pool excludes the country with the largest weight in the main estimates in Tables 2-4.
For brevity, the following ISO codes were used in the table: CHE for Switzerland, LUX for Luxembourg and, of course, USA for the United States.

TABLE E.8: Summary of SCM Estimates of the Impact of the Hartz Reforms on Unemployment Rate in Germany

Panel A: Main estimates	All	Men	Women		All	Men	Women	Actual
APEMR	0.04	0.07	0.02		Synthetic Germany			Germany
Pre-intervention RMSPE	0.43	0.59	0.22	Population weighted distance	8.48	8.72	8.63	8.74
P-value: Post-pre RMSPE ratio	0.04	0.37	0	Education expenditure	4.66	4.9	5.19	4.3
post/pre RMSPE ratio rank	2	11	1	Domestic credit by banking sector	102.64	137.08	82.32	130.24
Donour probability	0.07	0.41	0.04	Openness at 2005 constant prices	53.08	61.46	51.18	51.18
Estimated impact	-4.28	-0.08	-2.75	Consumption share of GDP	73.76	68.19	72.54	70.03
	Greece	France	Sweden	Government share of GDP	8.38	7.28	7.15	6.1
Top five countries	Sweden	Japan	Turkey	Investment share of GDP	22.5	23.7	19.56	23.02
with the largest weights	Portugal	Estonia	Italy	Per capita GDP	9.54	9.72	9.41	9.98
	Japan	Belgium	Greece	Population 65 and older	15.61	14.67	13.23	15.77
	Turkey	Sweden	Finland	Years of schooling	9.26	9.96	8.74	9.64

Panel B: Alternative set of predictors	All	Men	Women
APEMR	0.05	0.07	0.02
Pre-intervention RMSPE	0.47	0.58	0.25
P-value: Post-pre RMSPE ratio	0.26	0.48	0
post/pre RMSPE ratio rank	8	14	1
Donour probability	0.3	0.52	0.04
Estimated impact	-0.27	0.03	-1.69
	France	France	Sweden
Top five countries	Japan	Japan	France
with the largest weights	Sweden	Belgium	Turkey
	Turkey	Estonia	Israel
	Estonia	Israel	Estonia

Panel C: Alternative donor pool	All	Men	Women
APEMR	0.05	0.07	0.04
Pre-intervention RMSPE	0.45	0.63	0.43
P-value: Post-pre RMSPE ratio	0.19	0.42	0.15
post/pre RMSPE ratio rank	6	12	5
Donour probability	0.23	0.46	0.19
Estimated impact	-0.1	-0.23	-0.52
	Sweden	Belgium	France
Top five countries	Belgium	Estonia	USA
with the largest weights	Japan	Japan	Japan
	France	Canada	CHE
	Israel	Sweden	Finland

The data period is 1991-2015.
Intervention (beginning of the Hartz reforms) is in 2003. Thus, the pre-intervention period is 1991-2002.
APEMR = absolute prediction error to mean ratio, Estimated impact = annual average post-intervention gap between actual and synthetic.
Population weighted distance and per capita GDP are in logarithms.
Panel A has the same specification as Tables 2-4.
Panel B has the same specification as Table 5.
Panel C has the same specification as Table 6. The excluded countries are: Greece (all), France (men) and Sweden (women).
The estimates show that while the Hartz reforms had a significance effect on overall unemployment rate and women's unemployment rate show in Panel A, subsequently in Panels B and C, they turn out to be non-robust.

Bibliography

Abadie, A. and J. Gardeazabal (2003). "The Economic Costs of Conflict: A Case Control Study for the Basque Country". In: *American Economic Review* 93.1, pp. 113–132.

Abadie, A., A. Diamond, and J. Hainmueller (2010). "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program". In: *Journal of the American Statistical Association* 105.490, pp. 493–505. ISSN: 0162-1459.

Abadie, A. and A. Diamond (2015). "Comparative Politics and the Synthetic Control Method". In: *American Journal of Political Sciences* 59.2, pp. 495–510.

Acemoglu, D., S. Johnson, and J. Robinson (2001). "The Colonial Origins of Comparative Development: An Empirical Investigation". In: *American Economic Review* 91.5, pp. 1369–1401.

Acemoglu, D., P. Antràs, and E. Helpman (2007). "Contracts and Technology Adoption". In: *American Economic Review* 97.3, pp. 916–943.

Achterbosch, T. *et al.* (2009). "Measure the measure: the impact of differences in pesticide MRLs on Chilean fruit exports to the EU". In:

Agricultural Outlook 2015-2024. Parais: OECD Publishing. ISBN: 978-92-64-23203-7.

Akyol, M. and J. C. Beghin, eds. (2005). *Global agricultural trade and developing countries*. Trade and development series. Washington, D.C.: World Bank. ISBN: 9780821358634.

Akyol, M., M. Neugart, and S. Pichler (2013). "Were the Hartz Reforms Responsible for the Improved Performance of the German Labour Market?" In: *Economic Affairs* 33.1, p. 34.47.

Anders, S. M. and J. A. Caswell (2009). "Standards as Barriers Versus Standards as Catalysts: Assessing the Impact of HACCP Implementation on U.S. Seafood Imports". In: *American Journal of Agricultural Economics* 91.2, pp. 310–321. ISSN: 00029092.

Anderson, J. E. and E.v. van Wincoop (2003). "Gravity with Gravitas: A Solution to the Border Puzzle". In: *American Economic Review* 93.1, pp. 170–192.

Baier, S. and J. H. Bergstrand (2010). "Approximating General Equilibrium Impacts of Trade Liberalizations using the Gravity Equation". In: *The Gravity Model in International Trade: Advances and Applications*. Ed. by P. van Bergeijk and S. Brakman. Cambridge University Press, pp. 88–134.

Balogh, P. *et al.* (2016). "Consumer willingness to pay for traditional food products". In: *Food Policy* 61, pp. 176–184.

Batte, M. T. *et al.* (2007). "Putting their money where their mouths are: Consumer willingness to pay for multi-ingredient, processed organic food products". In: *Food Policy* 32.2, pp. 145–159.

Beghin, J. C., M. Maertens, and J. Swinnen (2015). "Nontariff Measures and Standards in Trade and Global Value Chains". In: *Annual Review of Resource Economics* 7.1, pp. 425–450. ISSN: 1941-1340.

Bénassy-Quéré, A., M. Coupet, and T. Mayer (2007). "Institutional Determinants of Foreign Direct Investment". In: *The World Economy* 30.5, pp. 764–782. ISSN: 0378-5920.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). "How Much Should We Trust Differences-in-Differences Estimates?" In: *Quarterly Journal of Economics* 119.1, pp. 249–275.

Bhosale, J. (2017). "Indian onion exports rise after minimum export price removal". In: *The Economic Times*.

Bloom, N., R. Sadun, and J. van Reenen (2012). "The Organization of Firms Across Countries". In: *The Quarterly Journal of Economics* 127.4, pp. 1663–1705. ISSN: 0033-5533.

Bohn, S. *et al.* (2014). "Did the 2007 Legal Arizona Workers Act Reduce the State's Unauthorized Immigrant Population?" In: *Review of Economics and Statistics* 96.2, pp. 258–269. ISSN: 0034-6535.

Bouvard, F. *et al.* (2013). "Have the Hartz reforms shaped the German labour market?" In: *Trésor-Economics* 110.

Buchmueller, T., J. DiNardo, and R. G. Valletta (2011). "The Effect of an Employer Health Insurance Mandate on Health Insurance Coverage and the Demand for Labor: Evidence from Hawaii". In: *American Economic Journal: Economic Policy* 3.4, pp. 25–51.

Burda, M. (2006). "Factor Reallocation in Eastern Germany after Reunification". In: *American Economic Review* 96.2, pp. 368–374.

Burda, M. and S. Seele (2016). "No role of the Hartz reforms? Demand and supply factors in the German labor market, 1993-2014". In: *SFB 649 Discussion Paper* 10.

Caliendo, L. and E. Rossi-Hansberg (2012). "The Impact of Trade on Organization and Productivity". In: *The Quarterly Journal of Economics* 127.3, pp. 1393–1467. ISSN: 0033-5533.

Cameron, A. C. and P. K. Trivedi (2010). *Microeconometrics using Stata*. Lakeway Drive: Stata Press. ISBN: 978-1-59718-073-3.

Chege, C. G.K., C. I.M. Andersson, and M. Qaim (2015). "Impacts of Supermarkets on Farm Household Nutrition in Kenya". In: *World Development* 72, pp. 394–407. ISSN: 0305750X.

Chen, Ch. (2008). "Measuring the Effect of Food Safety Standards on China's Agricultural Exports". In: *Review of World Economics* 144.1, pp. 83–106.

Chen, M., T. Otsuki, and J. S. Wilson (2006). "Do Standards Matter for Export Success?" In: *World Bank Policy Research Working Paper* 3809.

Chiputwa, B., D. Spielman, and M. Qaim (2015). "Food Standards, Certification, and Poverty among Coffee Farmers in Uganda". In: *World Development* 66, pp. 400–412. ISSN: 0305750X.

Colen, L., M. Maertens, and J. Swinnen (2012). "Private Standards, Trade and Poverty: GlobalGAP and Horticultural Employment in Senegal". In: *The World Economy* 35.8, pp. 1073–1088. ISSN: 03785920.

Crimmann, A. *et al.* (2009). "Forschungsbericht zum Thema Arbeitnehmerüberlassung". In: *Forschungsbericht Arbeitsmarkt* 397.

Czubala, W., B. Shepherd, and J. S. Wilson (2009). "Help or Hindrance? The Impact of Harmonised Standards on African Exports". In: *Journal of African Economies* 18.5, pp. 711–744. ISSN: 0963-8024.

Dauser, T. (2009). "Arbeitnehmer zweiter Klasse? Wild-West-Methoden in der Leiharbeitsindustrie". In: *Zeitarbeit*. Ed. by M.-O. Schwaab and A. Durian. Wiesbaden: Gabler, pp. 172–192. ISBN: 978-3-8349-1277-0.

David, P. A. and S. Greenstein (1990). "The Economics Of Compatibility Standards: An Introduction To Recent Research 1". In: *Economics of Innovation and New Technology* 1.1-2, pp. 3–41. ISSN: 1043-8599.

de la Mata, T. and C. Llano (2013). "Social networks and trade of services: modelling interregional flows with spatial and network autocorrelation effects". In: *Journal of Geographical Systems* 15.3, pp. 319–367. ISSN: 1435-5930.

Disdier, A. C. and S. Marette (2010). "The Combination of Gravity and Welfare Approaches for Evaluating Nontariff Measures". In: *American Journal of Agricultural Economics* 92.3, pp. 713–726. ISSN: 00029092.

Djankov, S., C. Freund, and C. Pham (2010). "Trading on Time". In: *The Review of Economics and Statistics* 92.1, pp. 166–173.

Dollar, D. and A. Kraay (2003). "Institutions, trade, and growth". In: *Journal of Monetary Economics* 50.1, pp. 133–162. ISSN: 03043932.

– (2004). "Trade, growth, and poverty". In: *The Economic Journal* 114, pp. 22–49. ISSN: 00130133.

Donald, S. and K. Lang (2007). "Inference with Difference-in-Differences and Other Panel Data". In: *Review of Economics and Statistics* 89.2, pp. 221–233. ISSN: 0034-6535.

Doyle, E. and I. Martinez-Zarzoso (2011). "Productivity, Trade, and Institutional Quality: A Panel Analysis". In: *Southern Economic Journal* 77.3, pp. 726–752.

Drogué, S. and F. DeMaria (2012). "Comparing apples with pears. How differences in pesticide residues regulatoins impact trade?" In: *Working Paper MOISA*.

Dustmann, C. *et al.* (2014). "From Sick Man of Europe to Economic Superstar: Germany's Resurgent Economy". In: *Journal of Economic Perspectives* 28.1, pp. 167–188. ISSN: 0895-3309.

Eaton, D. L. and J. D. Groopman, eds. (1994). *The Toxicology of Aflatoxins: Human Health, Veterinary, and Agricltural Significance*. San Diego, Carlifornia: Academic Press. ISBN: 0-12-228255-8.

Ebbinghaus, B. and W. Eichhorst (2006). "Employment Regulation and Labor Market Policy in Germany, 1991-2005". In: *IZA Working Paper Series* 2505.

Economist (1999). "The sick man of Europe". In: *Economist*. URL: http://www.economist.com/node/209559.

– (2013). "Dissecting the miracle". In: *The Economist*. URL: http://www.economist.com/news/special-report/21579145-ingredients-german-economic-success-are-more-complex-they-seem-dissecting.

– (2017). "In retreat: The multinational company is in trouble: Global firms are surprisingly vulnerable to attack". In: *The Economist* 28.01.2017. URL: http://www.economist.com/news/leaders/21715660-global-firms-are-surprisingly-vulnerable-attack-multinational-company-trouble.

Eichengreen, B. and D. A. Irwin (1998). "The Role of History in Bilateral Trade Flows". In: *The Regionalization of the World Economy*. Ed. by J. Frankel. University of Chicago Press, pp. 33–62. ISBN: 0-226-25995-1.

European Central Bank (2012). "Euro area labour markets and the crisis - Structural Issues Report October 2012". In:

European Commission (2017a). *European Food Safety*. URL: http://ec.europa.eu/food/safety\textunderscoreen.

– (2017b). *The European Commission for Health and Food Safety*. URL: http://ec.europa.eu/dgs/health\textunderscorefood-safety/index\textunderscoreen.htm.

FAO and WHO (2016). *Codex Alimentarius*. URL: http://www.fao.org/fao-who-codexalimentarius/en/.

Federal Institute for Population Research (14.08.2013). *Erstmals seit 1974 mehr Erwerbstätige als Rentner unter den 60- bis 65-Jährigen*. Wiesbaden.

Ferguson, S. (2009). "Exporting and Quality Complementarity with Heterogeneous Firms". In: *unpublished manuscript*.

Fernandes, A. M., E. Ferro, and J. S. Wilson (2015). "Product Standards and Firms' Export Decisions". In: *World Bank Policy Research Working Paper* 7315.

Ferro, E., J. S. Wilson, and T. Otsuki (2013). "The Effect of Product Standards on Agricultural Exports from Developing Countries". In: *Policy Research Working Paper* 6518.

– (2015). "The effect of product standards on agricultural exports". In: *Food Policy* 50, pp. 68–79.

Figueiredo, E., L. R. Lima, and G. Schaur (2014). "Robust Estimation of Gravity Equations and the WTO Impact on Trade Inequality: CESifo Conference on Estimation of Gravity Models of Bilateral Trade". In: *CESifo Conference*.

Francois, J. and M. Manchin (2013). "Institutions, Infrastructure, and Trade". In: *World Development* 46, pp. 165–175. ISSN: 0305750X.

Frankel, J. and D. Romer (1999). "Does Trade Cause Growth?" In: *American Economic Review* 89.3, pp. 379–399.

Fulponi, L. (2006). "Private voluntary standards in the food system: The perspective of major food retailers in OECD countries". In: *Food Policy* 31.1, pp. 1–13.

Garicano, L. and E. Rossi-Hansberg (2015). "Knowledge-Based Hierarchies: Using Organizations to Understand the Economy". In: *Annual Review of Economics* 7.1, pp. 1–30. ISSN: 1941-1383.

Gaskarth, G. (2014). "The Hartz Reforms ... and their lessons for the UK". In: *Centre for Policy Studies*.

GlobalGAP. *GlobalGAP Annual Reports*. Ed. by GlobalGAP. Cologne.

Goecke, H. *et al.* (2013). "Zehn Jahre Agenda 2010 – Eine empirische Bestandsaufnahme ihrer Wirkungen". In: *IW Policy Paper* 7.

Graffham, A., E. Karehu, and J. MacGregor (2007). *Impact of Eurep GAP on small-scale vegetable growers in Kenya*. Vol. 6. London: International Institute for Environment and Development.

Guimaraes, P. and P. Portugal (2010). "A simple feasible procedure to fit models with high-dimensional fixed effects". In: *The Stata Journal* 10.4, pp. 628–649.

Hall, R. and C. Jones (1999). "Why do some countries produce so much more output per worker than others". In: *The Quarterly Journal of Economics*, pp. 83–116. ISSN: 0033-5533.

Hansen, H. and N. Trifković (2014). "Food Standards are Good – For Middle-Class Farmers". In: *World Development* 56, pp. 226–242. ISSN: 0305750X.

Head, K. and T. Mayer (2014). "Gravity Equations: Workhorse, Toolkit, and Cookbook". In: *Handbook of International Economics*. Vol. 4, pp. 131–195.

Heckman, J. J. (1979). "Sample Selection Bias as a Specification Error". In: *Econometrica* 47.1, p. 153.

Helpman, E. (2011). *Understanding Global Trade*. The Belknap Press of Harvard University Press. ISBN: 978-0-674-06078-4.

Helpman, E., M. J. Melitz, and Y. Rubinstein (2008). "Estimating Trade Flows: Trading Partners and Trading Volumes". In: *Quarterly Journal of Economics* 73.2.

Henson, S. and W. Mitullah (2004). "Kenyan Exports of Nile Perch: The Impact of Food Safety Standards on an Export-Oriented Supply Chain". In: *World Bank Policy Research Working Paper* 3349.

Henson, S. and S. Jaffee (2008). "Understanding Developing Country Strategic Responses to the Enhancement of Food Safety Standards". In: *The World Economy* 31.4, pp. 548–568. ISSN: 0378-5920.

Henson, S. and J. Humphrey (2010). "Understanding the complexities of private standards in global agri-food chains as they impact developing countries". In: *The journal of development studies* 46.9, pp. 1628–1646. ISSN: 0022-0388.

Herzfeld, T., L. S. Drescher, and C. Grebitus (2011). "Cross-national adoption of private food quality standards". In: *Food Policy* 36.3, pp. 401–411.

Hoekman, B., ed. (2015). *The Global Trade Slowdown: A New Normal?* London: CEPR Press.

Hoekman, B. and A. Nicita (2011). "Trade Policy, Trade Costs, and Developing Country Trade". In: *World Development* 39.12, pp. 2069–2079. ISSN: 0305750X.

Hummels, D. and P. Klenow (2005). "The Variety and Quality of a Nation's Exports". In: *American Economic Review* 95.3, p. 704.723.

International featured standard (2016). Berlin. URL: https://www.ifs-certification.com/index.php/de/retailer/how-does-the-ifs-ensure-the-quality-of-your-products-de.

Jacobi, L. and J. Kluve (2006). "Before and After the Hartz Reforms: The Performance of Active Labour Market Policy in Germany". In: *IZA Discussion Paper Series*.

Jaffee, S. and S. Henson (2004). "Standards and Agro-Food Exports from Developing Countries: Rebalancing the Debate". In: *World Bank Policy Research Working Paper* 3348.

– (2005). "Agro-food Exports from Developing Countries: The Challenges Posed by Standards". In: *Global agricultural trade and developing countries*. Ed. by M. Akyol and J. C. Beghin. Trade and development series. Washington, D.C.: World Bank, pp. 91–114. ISBN: 9780821358634.

Kamal, F. and A. Sundaram (2016). "Buyer–seller relationships in international trade: Do your neighbors matter?" In: *Journal of International Economics* 102, pp. 128–140.

Kaufmann, D., A. Kraay, and M. Mastruzzi (2010). "The Worldwide Governance Indicators: Methodology and Analytical Issues". In: *World Bank Policy Research Working Paper* 5430.

Kettner, A. and M. Rebien (2007). "Hartz-IV-Reform: Impulse für den Arbeitsmarkt". In: *IAB Kurzbericht* 19.

Kiss, P. and P. Weingarten (2003). "Cost of compliance with the aquis communautaire in the Hungarian dairy sector". In: *Discussion Paper Institute of Agricultural Development in Central and Eastern Europe* 55.

Kleinwechter, U. and H. Grethe (2006). *Adoption of Eurepgap standards by mango exporters in Piura, Peru*.

Knuth, M. (2014). *Labour market reforms and the jobs miracle in Germany*. Visits and Publication Unit. ISBN: 978-92-830-2210-7.

Koenig, P., F. Mayneris, and S. Poncet (2010). "Local export spillovers in France". In: *European Economic Review* 54.4, pp. 622–641. ISSN: 00142921.

Koenker, R. (2005). *Quantile Regression*. Cambridge University Press. ISBN: 978-0-521-60827-5.

Krause, M.. and H. Uhlig (2012). "Transitions in the German labor market: Structure and crisis". In: *Journal of Monetary Economics* 59.1, pp. 64–79. ISSN: 03043932.

Krebs, T. and M. Scheffel (2013). "Macroeconomic Evaluation of Labor Market Reform in Germany". In: *IMF Economic Review* 61.4, pp. 664–701. ISSN: 2041-4161.

Krugman, P. (1979). "Increasing returns, monopolistic competition, and international trade". In: *Journal of International Economics* 9, pp. 469–479.

Latouche, K. and E. Chevassus-Lozza (2015). "Retailer Supply Chain and Market Access: Evidence From French Agri-food Firms Certified with Private Standards". In: *The World Economy* 38.8, pp. 1312–1334. ISSN: 03785920.

Launov, A. and K. Wälde (2013). "Estimating incentive and welfare effects nonstationary unemployment benefits". In: *International Economic Review* 54.4, pp. 1159–1198.

Levchenko, A. (2007). "Institutional Quality and International Trade". In: *Review of Economic Studies* 74.3, pp. 791–819.

Li, Y. and J. C. Beghin (2012). "A meta-analysis of estimates of the impact of technical barriers to trade". In: *Journal of Policy Modeling* 34.3, pp. 497–511. ISSN: 01618938.

Maertens, A. and C. B. Barrett (2013). "Measuring Social Networks' Effects on Agricultural Technology Adoption". In: *American Journal of Agricultural Economics* 95.2, pp. 353–359. ISSN: 00029092.

Maertens, M. and J. Swinnen (2009). "Trade, Standards, and Poverty: Evidence from Senegal". In: *World Development* 37.1, pp. 161–178. ISSN: 0305750X.

Mangelsdorf, A., A. Portugal-Perez, and J. S. Wilson (2012). "Food standards and exports: evidence for China". In: *World Trade Review* 11.03, pp. 507–526. ISSN: 1474-7456.

Martínez-Zarzoso, I. (2013). "The log of gravity revisited". In: *Applied Economics* 45.3, pp. 311–327. ISSN: 0003-6846.

– (2015). "The trade effects of foreign aid: an empirical perspective". In: *Handbook on the Economics of Foreign Aid*. Ed. by B. Mak Arvin and B. Lew. Cheltenham: Edward Elgar Publishing, pp. 179–196. ISBN: 978 1 78347 457 8.

Maskus, K. and T. Otsuki (2005). "The Cost of Compliance with Product Standards of Firms in Developing Countries: An Econometric Study". In: *World Bank Policy Research Working Paper* 3590.

Masood, A. and B. Brümmer (2014). "Impact of GlobalGAP Certification on EU Banana Imports: A Gravity Modeling Approach". In: *GlobalFood Disucssion Paper* 49.

Mayer, T. and G. Ottaviano (2007). "The Happy Few: The internationalisation of European firms: New facts based on firm-level evidence". In: *Bruegel Blueprint Series* 3.

McCaffrey, D. F. *et al.* (2012). "Review of Stata Routines for Fixed Effects Estimation in Normal Linear Models". In: *The Stata Journal* 12.3, pp. 406–432.

McClusky, J. J. (2007). "Public and Private Food Quality Standards: Recent Trends and Strategic Incentives". In: *Global supply chains, standards and the poor*. Ed. by J. Swinnen. Wallingford: CABI, pp. 19–25. ISBN: 9781845931858.

Melitz, M. (2003). "The impact of trade on intra-industry reallocations and aggregate industry productivity". In: *Econometrica* 71.6, p. 1695.1725.

Melo, O. *et al.* (2014). "Do Sanitary, Phytosanitary, and Quality-related Standards Affect International Trade? Evidence from Chilean Fruit Exports". In: *World Development* 54, pp. 350–359. ISSN: 0305750X.

Mideksa, T. (2013). "The economic impact of natural resources". In: *Journal of Environmental Economics and Management* 65.2, pp. 277–289.

Morici, P. (2017). "The 45 percent tariff: Trump must be realistic in negotiations with China". In: *The Washington Times* 22.01.2017. URL: http://www.washingtontimes.com/news/2017/jan/22/donald-trumps-45-percent-tariff-on-china/.

Muange, E. N., S. Schwarze, and M. Qaim (2014). "Social networks and farmer exposure to improved crop varieties in Tanzania". In: *GlobalFood Disucssion Paper* 45.

Munasib, A. and D. Roy (2011). "Sanitary and Phytosanitary Standards as Bridge to Cross". In: *IFPRI Discussion Paper* 01140.

Munasib, A. and D. Rickman (2015). "Regional economic impacts of the shale gas and tight oil boom: A synthetic control analysis". In: *Regional Science and Urban Economics* 50, pp. 1–17.

Nagl, W. and M. Weber (2014). "Unemployment compensation and unemployment duration before and after the German Hartz IV reform". In: *Ifo Working Paper Series* 186.

Neumark, D., Salas, J. M. I., and W. Wascher (2014). "Revisiting the Minimum Wage-Employment Debate: Throwing Out the Baby with the Bathwater?" In: *Industrial and Labor Relations Review* 67.2.5, pp. 608–648.

Nunn, N. (2007). "Relationship-specificity, incomplete contract, and the pattern of trade". In: *The Quarterly Journal of Economicy*, pp. 569–600.

Nunn, N. and D. Trefler (2013). *Domestic Institutions as a Source of COmparative Advantage*.

Ochel, W. (2005). "Hartz IV - Welfare to Work in Germany". In: *CESifo DICE Report* 3.2, pp. 18–25.

OECD (2006). "OECD Economic Surveys: Germany 2006". In: *OECD Publishing*.

– (2016a). *Harmonized Unemployment Rate*. URL: https://data.oecd.org/unemp/harmonised-unemployment-rate-hur.htm.

– (2016b). *Hours worked*. URL: https://data.oecd.org/emp/hours-worked.htm\#indicator-chart.

– (2016c). *Labor force participation rate*. URL: https://data.oecd.org/emp/labour-force-participation-rate.htm\#indicator-chart.

– (2016d). *Temporary Employment Indicator*. URL: https://data.oecd.org/emp/temporary-employment.htm\#indicator-chart.

Otsuki, T., J. S. Wilson, and K. Maskus (1999). "Quantifying the impact of technical barriers to trade: A framework for analysis". In: *World Bank Policy Research Working Paper*.

Otsuki, T., J. S. Wilson, and M. Sewadeh (2001). "Saving two in a billion: quantifying the trade effect of European food safety standards on African Exports". In: *Food Policy* 26, pp. 495–514.

Pomeranz, D. (2015). "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax". In: *American Economic Review* 105.8, pp. 2539–2569.

Pritchett, L. (2003). "Comment on: Institutions, trade, and growth". In: *Journal of Monetary Economics* 50.1, pp. 163–165. ISSN: 03043932.

Rauch, J. and A. Casella (2003). "Overcoming informational barriers". In: *The Economic Journal* 113, pp. 21–42. ISSN: 00130133.

Rinne, U. and K. Zimmermann (2013). "Is Germany the North Star of Labor Market Policy?" In: *IMF Economic Review* 61.4, pp. 702–729.

Rodrik, D. (2012). *The Globalization Paradox: Democracy and the Future fo the World Economy*. Norton & Company. ISBN: 978-0-393-34128-7.

Roy, D. (2013). "Aflatoxins - finding solutions for improved food safety: Trade impacts of aflatoxin standards". In:

Rubin, D. B. (2001). "Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation". In: *Health Services and Outcomes Research Methodology* 1, pp. 169–188.

Santos Silva, J.M.C. and S. Tenreyro (2006). "The Log of Gravity". In: *The Review of Economics and Statistics* 88.4, pp. 641–658.

– (2011). "Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator". In: *Economics Letters* 112.2, pp. 220–222. ISSN: 01651765.

Schäfer, H. (2003). "Reform der Arbeitslosenversicherung - Ökonomische Aspekte einer politische Debatte". In: *IW-Positionen*.

Scheepers, S., A. Jooste, and Z. G. Alemu (2007). "Quantifying the impact of phytosanitry standards with spesific reference to MRLs on the trade flow of South African avocados to the EU". In:

Schieritz, M. (2017). "Gegenangriff: Die EU hat begonnen, sich auf einen Handelskrieg gegen die USA vorzubereiten". In: *Die Zeit* 7. URL: http://www.zeit.de/2017/07/eu-donald-trump-usa-handelskrieg-freihandel.

Schuster, M. and M. Maertens (2013). "Do private standards create exclusive supply chains? New evidence from the Peruvian asparagus export sector". In: *Food Policy* 43, pp. 291–305.

– (2015). "The Impact of Private Food Standards on Developing Countries' Export Performance: An Analysis of Asparagus Firms in Peru". In: *World Development* 66, pp. 208–221. ISSN: 0305750X.

Seeleib-Kaiser, M. and T. Fleckenstein (2007). "Discourse, Learning and Welfare State Change: The Case of German Labour Market Reforms". In: *Social and Policy Administration* 41.5, pp. 427–488.

Shepherd, Ben and Norbert L.W. Wilson (2013). "Product standards and developing country agricultural exports: The case of the European Union". In: *Food Policy* 42, pp. 1–10.

Shingal, A., M. Ehrich, and L. Foletti (2016). "Re-estimating the effects of stricter standards on trade: endogeneity matters". In: *unpublished manuscript*.

Snower, D., Merkl, C. (2006). "The Caring Hand that Cripples: The East German Labor Market after Reunification". In: *American Economic Review#* 96.2, pp. 375–382.

Spermann, A. (2011). "The New Role of Temporary Agency Work in Germany". In: *IZA Discussion Paper Series* 6180.

Swinnen, J., ed. (2007). *Global supply chains, standards and the poor: How to globalization of food systems and standard affects rural development and poverty*. Wallingford: CABI. ISBN: 9781845931858.

Swinnen, J. *et al.* (2015). *Quality standards, value chains, and international development*. New York: Cambridge University Press. ISBN: 978-1-107-02591-2.

Vigani, M., V. Raimondi, and A. Olper (2012). "International trade and endogenous standards: the case of GMO regulations". In: *World Trade Review* 11.03, pp. 415–437. ISSN: 1474-7456.

Wilson, J. S., T. Otsuki, and B. Majumdsar (2003). "Balancing food safety and risk: do drug residue limits affect international trade in beef?" In: *Journal of International Trade & Economic Development* 12.4, pp. 377–402. ISSN: 09638199.

Wilson, John S. and T. Otsuki (2004). "To spray or not to spray: pesticides, banana exports, and food safety". In: *Food Policy* 29.2, pp. 131–146.

Winchester, N. *et al.* (2012). "The Impact of Regulatory Heterogeneity on Agri-food Trade". In: *World Economy* 35.8, pp. 973–993. ISSN: 03785920.

Wooldridge, J. (2010). *Econometric Analysis of cross section and panel data*. 2nd ed. Cambridge, MA: MIT Press. ISBN: 978-0-262-23258-6.

WTO (2013). *Aid for trade and value chains in agrifood*.

Xiong, B. and J. Beghin (2012). "Does European aflatoxin regulation hurt groundnut exporters from Africa?" In: *European Review of Agricultural Economics* 39.4, pp. 589–609. ISSN: 0165-1587.

Xiong, B. and J. C. Beghin (2013). "Disentangling demand-enhancing and trade-cost effects of maximum residue regulations". In: *Iowa state university Working Paper* 11019.

– (2014). "Disentangling demand-enhancing and trade-costs effects of maximum residue regulations". In: *Economic Inquiry* 52.3, pp. 1190–1203. ISSN: 00952583.

Yu, M. (2010). "Trade, democracy, and the gravity equation". In: *Journal of Development Economics* 91.2, pp. 289–300. ISSN: 03043878.

Yue, N. *et al.* (2010). "An empirical analysis of the impact of EU's new food safety standards on china's tea export". In: *International Journal of Food Science & Technology* 45.4, pp. 745–750. ISSN: 09505423.

Zajonz, D. (2016). ""Ich muss Angela Merkel einen Großteil der Schuld geben": Der Harvard-Ökonom Dani Rodrik warnt vor den Exzessen der Globalisierung. Er kritisiert auch Deutschlands Rolle im europäischen Integrationsprozess." In: *Die Zeit* 22.08.2016. URL: http://www.zeit.de/wirtschaft/2016-08/dani-rodrik-eu-angela-merkel-globalisierung-probleme/komplettansicht.

Eidesstattliche Erklärung

Versicherung gemäß §14 Prüfungs- und Studienordnung für den Promotionsstudiengang Wirtschaftswissenschaften

1. Die Gelegenheit zum vorliegenden Promotionsvorhaben ist mir nicht kommerziell vermittelt worden. Insbesondere habe ich keine Organisation eingeschaltet, die gegen Entgelt Betreuerinnen und Betreuer für die Anfertigung von Dissertationen sucht oder die mir obliegenden Pfichten hinsichtlich der Prüfungsleistungen für mich ganz oder teilweise erledigt.
2. Ich versichere, dass ich die eingereichte Dissertation "Essays on effects of policy interventions in the realm of food standards, trade, and the German labour market " selbstständig und ohne unerlaubte Hilfsmittel verfasst habe; fremde Hilfe habe ich dazu weder unentgeltlich noch entgeltlich entgegengenommen und werde dies auch zukünftig so halten. Anderer als der von mir angegebenen Hilfsmittel und Schriften habe ich mich nicht bedient. Alle wörtlich oder sinngemäß den Schriften anderer Autoren entnommenen Stellen habe ich kenntlich gemacht.
3. Die Richtlinien zur Sicherung der guten wissenschaftlichen Praxis an der Universität Göttingen werden von mir beachtet.
4. Eine entsprechende Promotion wurde an keiner anderen Hochschule im In- oder Ausland beantragt; die eingereichte Dissertation oder Teile von ihr wurden nicht für ein anderes Promotionsvorhaben verwendet.
5. Des Weiteren ist mir bekannt, dass Unwahrhaftigkeiten hinsichtlich der vorstehenden Erklärung die Zulassung zur Promotion ausschließen bzw. später zum Verfahrensabbruch oder zur Rücknahme des erlangten Titels berechtigen.

Signed: _____

Date: _____