

Global Economic Integration, Technology Diffusion and Climate Change

Inaugural-Dissertation

zur Erlangung des akademischen Grades eines Doktors
der Wirtschafts- und Sozialwissenschaften
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Diplom-Wirtschaftsingenieur (Elektrotechnik)

Michael (Georg) Hübler

aus Fulda (Hessen).

Kiel, 2009

Gedruckt mit Genehmigung der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

Dekan: Professor Dr. Thomas Lux
Erstberichterstattender: Professor Dr. Johannes Bröcker
Zweitberichterstattender: Professor Dr. Till Requate

Tag der Abgabe der Arbeit: 26. Oktober 2009

Tag der mündlichen Prüfung: 12. April 2010

Preface

Global economic integration has the potential to interconnect humans across the globe and to increase their wealth. It can occur in a number of ways such as trade, foreign direct investment (FDI) and offshore outsourcing, which are central issues in the following chapters, especially with respect to developing countries. The question is in how far global integration can have positive and negative side effects.

Technology diffusion is one particular positive side effect of global economic integration. It has the potential of improving productivities of developing countries and hence the living standards of humans.

Climate change is one particular negative side effect of global economic integration. Global economic integration potentially increases energy use and carbon emissions, given it has a positive impact on economic growth. This in turn increases the detrimental impacts of climate change on humans and nature. While the industrialized countries are the main sources of climate change, the developing countries will be hit most severely by climate change.

However, one solution might be the interconnection of the three issues. *Global economic integration* may enhance global welfare by fostering economic growth and mitigating *climate change* via *technology diffusion* at the same time. The following chapters try to shed some light on this threefold interconnection by using static-theoretical, dynamic-theoretical, econometric, numerical simulation and projection methods.

In the light of *global economic integration*, the first chapter deals with offshore outsourcing based on a stylized, static, microeconomic model. It assumes Cournot competition between an intermediate good supplier in a high-cost, high-productivity region, denoted by the West, and an intermediate good supplier in a low-cost, low-productivity region, denoted by the East. It examines the impact of changes in marginal production costs in the East and the West on the allocation of intermediate good production between East and West. It asks the questions in which cases cost changes in one of the regions likely reduce intermediate good production in both regions. It turns out that higher production costs in one region indeed reduce intermediate good production in both regions under certain conditions, since the intermediate good input factor can be replaced by another input factor. Moreover, the sensitivity of outsourcing activities to production cost changes is highest when the interregional cost differential is smallest.

Also in the light of *global economic integration*, the second chapter deals with international capital mobility based on a stylized, dynamic, macroeconomic model. It assumes perfect capital mobility between a high-income, high-productivity region, denoted by the North, and a low-income, low-productivity region, denoted by the South. Additionally, in the light of *technology diffusion*, the second chapter introduces dynamics of technological catching up of the South to the technology frontier given by the North. It examines the positive impact of capital mobility and technology diffusion on the development of the South with respect to aggregate production. It asks the question, in which cases this positive development effect likely fails. It turns out that the South can fall behind in terms of technologies or get trapped in a situation in which it is unable to attract foreign capital and embodied technologies if it is too far away from the technology frontier and if its absorptive capacity is too low. The paper reconciles the view that technological catching up is stronger the larger the technology gap with the alternative view that technological catching up is strongest at a medium technology gap.

In the light of *global economic integration*, the third chapter¹ deals with FDI, international trade and foreign aid with respect to developing countries based on state of the art econometric methods. Moreover, in the light of *technology diffusion*, the third chapter examines the potential of FDI, trade and aid for transferring technologies from industrialized countries to developing countries. In the light of *climate change*, it applies energy intensity as the dependent variable under scrutiny. It asks the question whether there is a significant impact of FDI inflows, imports and received aid on energy intensities of developing countries. It turns out that the hypothesis stating aggregate FDI inflows reduce energy intensities of developing countries is not confirmed in a robust significant way. Rather, foreign development aid seems to reduce energy intensities of recipient countries.

Furthermore, in the light of *global economic integration*, the fourth chapter deals with a multi-region, multi-sector computable general equilibrium (CGE) model including inter-national trade and international capital mobility based on the GTAP 7 data set. The model also distinguishes North and South, and additionally China, which is known to be highly integrated into the global economy. In the light of *technology diffusion*, it transfers the stylized approach for technology diffusion of

¹ The chapter is joint work with Andreas Keller. While Andreas Keller had some emphasize on data work and on elaborating the economic background, I had some emphasize on methodological and technical work. Nevertheless, we worked jointly and equally on all relevant aspects of the article.

chapter two into a CGE framework. It then exploits the existing empirical evidence on general and energy specific technology diffusion via FDI and trade for calibrating the model. In the light of *climate change*, the model at first calculates carbon contents of traded commodities. At second, it examines carbon based border tax adjustment under a contraction and convergence climate policy regime focusing on China as the global major CO₂ emitter. It asks the question whether carbon based border tax adjustment is an effective policy measure for emission reductions. It turns out that the developing countries will benefit when China joins the contraction and convergence regime, while the industrialized countries will be almost unaffected. When China does not join the regime and instead a carbon content based border tax is imposed, the industrialized countries will significantly benefit, while China will be significantly worse off. The effect of the border tax adjustment on the global carbon price and on global emissions seems negligible.

Finally, in the light of *climate change*, the fifth chapter² specifically deals with possible future health risks and costs of climate change in Germany based on future climate simulations. It uses temperature data gained from the high-resolution REMO model for the end of the 21st century. It examines the impact of rising temperature on mortality, hospitalization costs and labor productivity. It asks the question whether Germany can expect a significant rise in health risks and costs due to climate change if we refuse to foster mitigation and adaptation sufficiently. It turns out that there could be an average increase in the number of heat induced casualties by a factor of more than 3. Heat related hospitalization costs could increase 6-fold not including the cost of ambulant treatment. The future heat load could also reduce work performance resulting in an estimated output loss of between 0.1% and 0.5% of GDP.

Kiel, 23 October 2009

Michael Hübler

² This chapter is joint work with Gernot Klepper and Sonja Peterson. While Gernot Klepper and Sonja Peterson were my supervisors when working on the study “Kosten des Klimawandels” and the technical paper “Costs of Climate Change”, I did the major work covering all relevant aspects of the article.

Acknowledgements

The Kiel Institute for the World Economy provided a very inspiring and fruitful research environment via its internal and external research networks, guest researchers and guest lecturers within the Advanced Studies Program, comprehensive access to literature and data, and beautiful buildings at a beautiful location. The Kiel University provided very profound and helpful knowledge within the Doctoral Program Quantitative Economics.

I thank and praise God. I thank my supervisors and colleagues at Kiel University, at the Kiel Institute for the World Economy and at Darmstadt University of Technology. I thank my parents and grandparents, aunts and uncles and my friends and neighbours. I also thank my music teachers and my teachers at school. I have learnt a lot from all of them, and I would achieve nothing on my own.

Contents

Chapter 1:

The Labor Market Effects of Outsourcing Parts and Components

Part 1: A Simple Model with Cournot Competition

1	Introduction	2
2	A Brief Review of the Outsourcing Literature	4
3	The Model Structure	6
4	Cournot Competition	9
5	Allocation of Intermediate Good Production.....	11
6	Allocation of Production Factors	14
7	Transport Costs	16
8	Graphical Representations	17
9	Interpretation and Caveats	21
10	Conclusions.....	24
11	References.....	25

Part 2: Adverse Cournot Competition

1	Introduction.....	29
2	Model setup.....	29
3	Second order condition for a profit maximum.....	32
4	Condition for adverse Cournot competition	33
5	Errata.....	36
6	Conclusions.....	36
7	References.....	37

Chapter 2:

Avoiding the Trap: The Dynamic Interaction of North-South

Capital Mobility and Technology Diffusion

1	Introduction.....	52
2	Literature Background	54
3	The Nelson and Phelps Theory in the Context of International Capital Mobility	56
	3.1 A Re-Interpretation of the Nelson and Phelps Theory	58
	3.2 Innovation in the South.....	60
	3.3 International Factor Allocation in the Steady State.....	64
4	Endogenous International Capital Mobility.....	64
	4.1 The Model.....	65
	4.2 Interpretation.....	67
5	Discussion	71
6	Conclusion	74
7	References.....	76
8	Appendix	78

Chapter 3:
Energy Savings via FDI?
Empirical Evidence from Developing Countries

1	Introduction.....	83
2	The influence of FDI inflows on energy use	84
	2.1 Theoretical background	85
	2.2 Empirical literature on energy saving technology transfer via FDI	89
3	Panel data analysis	91
	3.1 Derivation of the empirical model.....	92
	3.2 The estimating equations	95
	3.3 Regression results	98
4	Robustness checks	102
	4.1 M&A correction of FDI.....	102
	4.2 Sectoral data for the United States as a source country.....	103
5	Conclusion	104
6	References.....	106
7	Appendix.....	109

Chapter 4:
Can Carbon Based Import Tariffs Effectively Reduce Carbon Emissions?

1	Introduction.....	120
2	The three region model	123
3	Carbon content of trade.....	125
4	Border tax adjustment	129
5	Conclusion	135
6	References.....	137
7	Appendix.....	139

Chapter 5:
Costs of Climate Change:
The Effects of Rising Temperatures on Health and Productivity in Germany

1	Introduction.....	146
2	Heat scenarios for Germany	149
	2.1 Employed climate models and climate scenarios	149
	2.2 Heat scenarios	150
	2.3 Coldness scenarios.....	153
3	Temperature induced fatalities.....	155
	3.1 Heat induced mortality	155
	3.2 Coldness induced mortality	160
4	Economic costs of heat	163
	4.1 Heat induced hospitalization costs.....	164
	4.2 Heat induced production loss.....	167
5	Conclusion	172
6	Acknowledgements.....	173
7	References.....	174

Figures and Tables

Chapter 1:

Figure 1: Elasticity ε of intermediate good production in the West relative to the East (V_{WE}) with respect to relative production cost changes (c_{WE}), example with halflogarithmic scale, $\alpha = 0.66$	13
Figure 2: Absolute production allocation depending on an increasing western wage	17
Figure 3: Profit maximizing reaction functions.....	19
Figure 4: Relative factor allocation depending on an increasing western wage.....	20

Chapter 2:

Figure 1: Areas of convergence and divergence of technology growth rates in the North and the South dependent on the distance to frontier	70
Figure 2: Rate of technological progress in the South dependent on the distance to frontier for different income shares of foreign capital in the South	80

Chapter 3:

Table 1: Estimation results for specifications A, B and C.....	99
Table 2: Definition of variables.....	109
Figure A1: Trends of the variables in the simplified model with 20 countries, 1979-2003	110
Table A1: Unit root test statistics (Augmented Dickey Fuller) for the variables in the simplified model	110
Table A2: Johansen cointegration test of the variables in the simplified model.....	111
Table A3: Descriptive statistics of the sample of 60 developing countries in 1975-2004	111
Figure A2: Time trends of energy intensity 1975-2004 in four countries of the sample.....	112
Figure A3: Time trends of FDI in percent of GDP 1975-2004 in four countries of the sample.....	112

Figure A4: Indicators of the Chinese economy, 1975-2004.....	115
Table A4: Pair wise correlations of the explanatory variables	115
Table A5: Estimation results for specifications A, B and C with M&A corrected FDI.....	116
Table A6: Estimation results for the sectoral dataset with the United States as source country.....	117

Chapter 4:

Figure 1: Carbon intensities of products.....	127
Figure 2: Implicit carbon contents of exported products.....	128
Figure 3: Interregional carbon contents of trade.....	129
Figure 4: CO ₂ emissions under a worldwide carbon price (scenario +chi)	130
Figure 5: CO ₂ emissions under a worldwide carbon price excluding China (scenario -chi)	131
Figure 6: CO ₂ price under different policy scenarios	132
Figure 7: Accumulated, discounted welfare effects of policy scenarios with respect to reference scenario –chi.....	133
Table 1: Parameters	142
Table 2: Variables.....	143

Chapter 5:

Table 1: Estimated impacts of the European heat wave in 2003	147
Figure 1a: Additional number of days with (strong or extreme) heat stress, REMO experiment A1B (2071-2100) minus control run CTL (1971-2000) in Germany.....	151
Figure 2b: Additional number of days with (strong or extreme) heat stress, REMO experiment B1 (2071-2100) minus control run CTL (1971-2000) in Germany.....	152
Figure 3: Additional number of days per year with strong or extreme heat stress in Frankfurt/Main (Hessen), moving averages over 5 years.....	153

Figure 3: Reduction in the number of days per year with (light, moderate, strong or extreme) cold stress, control run CTL (1971-2000) minus REMO experiment A1B (2071-2100) in Germany	154
Figure 4: Number of heat induced deaths p. a. in Germany, 2071-2100 (exponential extrapolation for extreme heat stress).....	158
Figure 5: Distribution of additional heat induced deaths p. a. per 10 km times 10 km sectors in Germany, 2071-2100, A1B (exponential extrapolation for extreme heat stress)	159
Figure 6: Distribution of the reduction in coldness induced deaths p. a. per 10 km times 10 km sectors in Germany, 2071-2100, A1B	161
Figure 7: Comparison of the increase in heat related deaths and the reduction in coldness related deaths p. a. in Germany, 2071-2100, A1B, both computed in 10 km times 10 km resolution without base observations from the past.....	162
Figure 8: Hospitalization costs p. a. in Germany, 2071-2100 (exponential extrapolation for extreme heat stress).....	167
Figure 9: Heat related production loss in Germany p. a., 2071-2100 under two different impact assumptions.....	169
Figure 10: Heat related income losses p. a. and per capita, 2071-2100 across German federal states in 2004-€.....	171

Chapter 1:
The Labor Market Effects of Outsourcing Parts and Components
Part 1: A Simple Model with Cournot Competition
Part 2: Adverse Cournot Competition

Published in *Aussenwirtschaft*, University of St. Gallen, 63(2), 167-194
<http://www.siaw.unisg.ch/org/siaw/web.nsf/d638de4e02e667cac12568f0002661cf/479b573ee825c1adc1256a310023726c!OpenDocument>

Kiel Working Paper No. 1564
<http://www.ifw-kiel.de/pub/kap>

Chapter 1:

The Labor Market Effects of Outsourcing Parts and Components

Part 1: A Simple Model with Cournot Competition *

Michael Hübler ⁺

May 15, 2008

Abstract

The paper analyses a partial equilibrium outsourcing model with Cournot competition in intermediate good production. Final production is located in Western Europe, whereas the intermediate good can be manufactured by a Western (outsourcing) or Eastern European supplier (offshore outsourcing). The paper asks the question how changes in production costs, in particular wages, affect labour input in the two regions in the presence of Cournot competition. The main results are: higher production costs in one region reduce intermediate good production in both regions leading to a substitution effect between high- and low-skilled labour intensive inputs rather than between Eastern and Western low-skilled labour intensive inputs. The sensitivity of outsourcing activities to production cost changes is highest when the interregional cost differential is smallest.

JEL Classifications: D24, D43, F20, J31

Keywords: Offshoring, outsourcing, Cournot competition,
intermediate goods, high-skilled, low-skilled

* Special thanks to Günther Rehme; further thanks to Frank Bickenbach, Sonja Peterson, Andreas Keller, Peter Nunnenkamp, Edwin van der Werf, Henning Klodt and Min-Ching Chiou for helpful comments.

⁺ Kiel Institute for the World Economy, Düsternbrooker Weg 120, 24100 Kiel, Germany, E-mail: michael.huebler@ifw-kiel.de, Tel: +49-431-8814-401.

1 Introduction

The outsourcing phenomenon has been a frequently discussed topic in the United States since the late 20th century. The debate on the influence of globalisation versus technical progress on the increasing income differential between high-skilled and low-skilled labour is still going on. Outsourcing, or more precisely offshore outsourcing across borders, has also been intensively discussed in the context of the European Union Eastern enlargement. The question of main interest is how outsourcing activities, that attempt to exploit lower production costs in Central and Eastern European countries, affect employment in Western European countries. (Real wages in the East average around one-fifth of the respective wage levels in the former EU-15).¹ This is the so-called vertical outsourcing in the spirit of Feenstra and Hanson (1996).² This issue is especially important for the debate on unions' bargaining and minimum wages.

Sinn (2005) describes Germany as a Bazaar Economy importing and exporting large amounts of goods but adding a low production value. Indeed, the value added divided by the output value in the German industry declined from 40.2 % to 34 % between 1970 and 2003 (Sinn 2005). For example, the German Porsche Cayenne is actually to a large extent produced in Bratislava. Additionally, the European automotive industry can be characterised by the following developments (Meissner and Jürgens 2007): There are not only few automobile final producers, but also an agglomeration tendency towards oligopolies of “mega suppliers” exploiting scale effects and being powerful enough to bargain with the big final producers. In particular, in many cases there are only a few suppliers of a certain module that is

¹ According to World Bank purchasing power parity estimates, Boeri and Brücker (2000).

² Outsourcing means removing part of the production process from the own company and buying it from an external company instead. The outsourced part can be a product or a service, but not raw material (Kirkegaard 2005). Offshore outsourcing in the current discussion is a certain type of outsourcing, namely outsourcing to a foreign country. FDI is distinguished from outsourcing and pure portfolio investment by the ownership criterion: The investor has a certain degree of influence on the foreign investment receiving country (by holding a company share of at least 10 %). The OLI paradigm by Dunning (1981) distinguishes three motives for FDI: ownership advantage, location advantage and internalisation advantage. In that respect the expression multinational firm is closely related to FDI (for instance Markusen and Maskus 2001). There are two types of FDI: vertical and horizontal. Vertical FDI is similar to offshore outsourcing, since part of the production process is done by the affiliate and the resulting intermediate good is transferred back. An important factor driving vertical FDI is the benefit from lower production costs abroad. Horizontal FDI and horizontal multinationals have little in common with offshore outsourcing, since the main purpose is direct access to a foreign market instead of exporting to that market. According to the definition by Kirkegaard (2005) offshoring in general encompasses offshore outsourcing as well as vertical FDI and refers to imports of intermediate goods across borders.

used in final assembly. Furthermore, price pressure forces the suppliers to cut costs and to offer their intermediate components at the common price given by the rivals' performance and the existing technologies.

Referring to this development within the European automotive industry, this paper examines an easily tractable microeconomic partial equilibrium model. However, the model can also be applied to other industries with similar structures. The model assumes one final producer and two suppliers of intermediate components who are forced into Cournot competition, since the final producer would neither accept buying the components at a price higher than the common price, nor would he accept relying on one supplier only. This would make the final producer totally dependent on the single supplier. Instead there are two suppliers with oligopolistic power. We assume that one of them has moved intermediate production towards Eastern Europe, that is into a low-wage region. The other supplier keeps production in the Western European region that has higher wages, but also a better productivity level. The paper deals with the question in what ways and how sensitively the labour inputs of the suppliers react when production costs, in particular wages, change. Wages change exogenously according to unions' bargaining and policies such as minimum wages. When the supplier, who is affected by the production cost change, adjusts his output quantity, the rival will react immediately adjusting his output quantity as well, both aiming to maximise profits. The question is, whether this leads to a one to one substitution of unskilled labour between the East and the West. It turns out that the sensitivity of outsourcing activities to production cost changes is highest when the interregional cost differential is smallest. Moreover, it is shown that higher production costs in one region reduce intermediate good production in both regions leading to a substitution effect between high- and low-skilled labour intensive inputs rather than between Eastern and Western low-skilled labour intensive inputs.

The paper is organised as follows: Section 2 summarises the related theoretical and empirical literature on international outsourcing. Section 3 describes the model structure, section 4 introduces Cournot competition of intermediate goods suppliers, section 5 derives the allocation of intermediate good production and the elasticity of relative interregional production, and section 6 analyses the resulting allocation of production factors focussing on labour. Section 7 introduces transport costs, and section 8 provides graphical representations. Section 9 discusses implications and caveats, and section 10 concludes.

2 A Brief Review of the Outsourcing Literature

Offshore Outsourcing has first been investigated theoretically and empirically in the case of the United States. Later on, a mainly empirical literature strand has examined the labour market effects of offshore outsourcing in the context of the European Union Eastern enlargement. Not surprisingly, a variety of theoretical models has been set up in order to better understand the determinants and effects of outsourcing activities.

Concerning the United States, Feenstra and Hanson (1996a and 1996b) show that rising imports, reflecting the outsourcing of production activities, contributed to the decline of relative employment and wages of unskilled workers during the 1980s. When firms outsource low-skilled labour intensive activities to low-wage countries and import the produced intermediate goods, this will shift employment towards skilled labour within industries. Bhagwati, Panagariya and Srinivasan (2004) refer to US outsourcing at the beginning of the 21st century as another phenomenon: Trade in services at arm's length that does not require geographical proximity of the buyer and the seller. This view emphasises the role of information and communication technology opening the possibility of outsourcing call centres, software programming and data analysis to Asia connected via fast internet data transfer.

Turning to the European Union, Marin et al. (2002) indicate that the most dynamic and innovative segments of the German industry invest in Eastern Europe and that exploiting low Eastern wages is one motive for outsourcing activities. Marin (2006) finds empirically that falling trade costs, lower levels of corruption and improvements in the contracting environment in Eastern Europe influence the level of intra-firm imports from the East to Austria and Germany indicating more outsourcing to the East. Marin (2004) goes one step further, stating that German and Austrian firms carry out outsourcing activities towards the East in order to take advantage of the abundant high-skilled labour there. She finds high educational levels among employees and more workers engaged in R&D and engineering in Eastern affiliates compared to firms in Germany and Austria. Braconier and Ekholm (2001) find opposite results. Marin (2004) shows small job losses in the West due to outsourcing, because outsourcing helps Western firms to stay competitive in accordance with Konings and Murphy (2001). In contrast to these studies Becker et al. (2005) find a more substantial replacement of Western jobs by jobs in Eastern affiliates.

Kirkegaard (2005) identifies European companies supplying or receiving outsourced or offshored goods and services and consumers of the resulting final commodities as winners. On the other hand companies being unable to adapt to the outsourcing boom and workers laid off due to outsourcing are potential losers.

In summary, the empirical outcomes point to some negative redistribution and employment effects for low-skilled employees. However, there is no reason for an outsourcing hysteria concerning production and jobs moving to low wage countries rapidly, considering that outsourcing improves efficiency and competitiveness and therefore reduces commodity prices under competitive markets making consumers of these products better off.

Various models have been recently developed due to the public awareness of outsourcing and its consequences. There are models with a continuum of (intermediate) goods or production stages in the tradition of Feenstra and Hanson (1996a) and Grossman and Helpman (2002 and 2003), models extending the classic Heckscher-Ohlin theory, models based on standard production functions combined in multiple stages, and there are models based on outsourcing cost considerations. Within the group of models with a continuum of (intermediate) goods, Wang (2006) develops a model of choosing between vertical integration and outsourcing depending on cost differentials, transport costs and costs of searching for intermediate good trade. Kohler (2004) models the reaction of a multi-stage industry with outsourcing to changes of the final good price and fragmentation costs. Mitra and Ranjan (2005) extend the outsourcing and FDI literature to dynamic behaviour with externalities and firm heterogeneity. According to their results temporary shocks can have permanent effects, and most productive firms move abroad first. A key issue in the research of Grossman and Helpman (2005) is the view of outsourcing as an activity that requires a costly search for a partner (in the home country or in a foreign country). Grossman and Rossi-Hansberg (2006a) propose a new conceptual framework of the global production process focussing on tradable tasks. They show that in contrast to neoclassical trade theories (under certain conditions) all domestic parties can share the gains from improved offshoring opportunities. (For a comprehensive theoretical treatment of offshoring see Grossman and Rossi-Hansberg 2006b). A reformulation of the four basic theorems of the Heckscher-Ohlin theory allowing for offshoring (fragmentation) is provided by Baldwin and Robert-Nicoud (2007). Markusen (2005) applies modules from the existing trade theory to numerical analyses. Besides other

results, his simulations suggest welfare gains for the South and the global economy, while the outsourcing Northern country may lose if it is large. Munch and Skaksen (2005) combine CES and Cobb-Douglas functions in a multi-stage setup. They point out that outsourcing to abroad worsens the wages for unskilled-workers, whereas this effect cannot be expected from outsourcing within a country. Furthermore, Senses (2006) illustrates how the wage elasticity of low-skilled labour demand increases in heavily outsourcing industries and how a decline in the share of unskilled labour at home lowers the elasticity on the other hand. Bandyopadhyay and Wall (2005a) derive the optimal amount of outsourcing for a given immigration level, while the model by Bandyopadhyay and Wall (2005b) includes an oligopolistic export sector and a competitive import-competing sector and shows that an outsourcing tax can be justified under a minimum wage but not under flexible wages. Among the models applying cost function representations Egger and Egger (2004a) include multinationals firms' competition in quantities and price-cost margins. Egger and Egger (2007) analyse the trade-off between transport costs and lower production costs induced by outsourcing in a multi-stage decision chain on market entry and price setting in a spatial world. Bartel, Lach and Sicherman (2005) show how technological progress lowers firms' adjustment costs of outsourcing.

Summing up, a large variety of models has been developed to examine different aspects related to outsourcing, such as labour market effects, transport costs versus lower production costs, searching for an outsourcing partner, the relationship of outsourcing and migration and the role of technological progress. Accordingly, outsourcing is a phenomenon with many facets, some of them investigated and understood, others not yet understood leaving room for further research in different directions.

This paper deals with a facet of high public interest, not sufficiently covered theoretically: The reaction of (offshore) outsourcing in the spirit of Feenstra and Hanson (1996a and 1996b) to changes in production costs in the presence of Cournot competition of suppliers of parts and components.

3 The Model Structure

The partial equilibrium model consists of final good Y production located in the Western European country, and intermediate good X production located in the

Western country and in the Eastern country. The model can be interpreted referring to the European automotive industry as motivated in the introduction. Y is produced with a Cobb-Douglas technology using the inputs H and X :

$$Y = H^\alpha (X_W + X_E)^{1-\alpha}; \quad 0.5 < \alpha < 1 \quad (1)$$

The final good Y can be an automobile produced in Germany via a complex process with a sophisticated technology.³ The final good producer takes demand, for example for his car type, as given by the market. (The final good market form is not of importance for the considerations in this paper). H is a high-skilled labour and high-technology intensive input, which is available in the Western area only and cannot be outsourced. This is a conservative assumption contrasting with Marin (2004) finding outsourcing of high-skilled labour towards Eastern Europe, but not crucial for the interpretation of the outcomes. For simplicity the production factor H encompasses all processes involving highly educated employees such as design, engineering and management, the necessary high technology capital and firm specific knowledge.

The intermediate good X production process includes all activities demanding low-skilled labour like manual work plus usual capital input. X represents for instance interior automotive parts like dashboards and seats. X can be manufactured in Western (X_W) and Eastern Europe (X_E). Low-skilled labour is supplied in both regions. That means, there is no offshore outsourcing with the purpose of getting access to well educated workers in the East as described by Marin (2004), nor service outsourcing according to Bhagwati, Panagariya and Srinivasan (2004), but it is offshore outsourcing in the spirit of Feenstra and Hanson (1996a and 1996b). Intermediate good production can be moved to the East when production costs or wages are cheaper in the East, afterwards the manufactured intermediates are imported into the West. X_W and X_E are homogeneous goods and perfect substitutes so that they can be summed up to $X = X_W + X_E$. The assumption $\alpha > 0.5$ implies a higher income share for the high-skilled labour and high-technology intensive input located in the West (under perfect competition). This means that the share of the production process which is mobile via outsourcing is less than half of whole production in terms of revenues. The intuition is that the main part of production is not outsourced but

³ The model could also describe any other suitable industry, where outsourcing of intermediate goods production occurs, and where a small number of intermediate goods producers compete.

kept within the company in contrast to the example by Sinn (2005).⁴ $\alpha > 0.5$ is plausible, because H encompasses high-skilled labour as well as high tech capital. The Cobb-Douglas function implies the possibility of replacing part of input X , produced with a high amount of low-skilled labour and a standard technology by input H , produced by high-skilled workers and modern technologies, and vice versa. The Y producer minimises production costs C for a given output Y :

$$\begin{aligned} \min C &= w_H H + p_W X_W + p_E X_E \\ \text{s. t. } Y &= H^\alpha (X_W + X_E)^{1-\alpha} \end{aligned} \quad (2)$$

w_H is the price of high-skilled labour intensive input H , p_W and p_E are the prices of X , manufactured in the Western or Eastern region. The final good Y producer is a price taker; he sets the input factor amounts according to the factor prices w_H , p_W and p_E .

The intermediate good is manufactured by independent rivalling firms located in Western Europe or Eastern Europe, respectively. We call the former case outsourcing, the latter case offshore outsourcing. Intermediate good X production is represented by using Cobb-Douglas functions with the inputs capital K and low-skilled labour L with constant returns to scale:

$$X_i = A_i K_i^{\beta_i} L_i^{1-\beta_i}; \quad 0 < \beta_i < 1; \quad i = [W; E] \quad (3)$$

The production processes in the West and in the East differ in technologies A_i and in the real wages w_i . Capital is not mobile across borders. Furthermore, differences in the returns to capital investment r_i and in the elasticities of production β_i and $1-\beta_i$ are possible. Cost minimizing X manufacturing leads to the following marginal costs c_i (Varian 2001):

$$c_i = \beta_i^{-\beta_i} (1 - \beta_i)^{\beta_i - 1} A_i^{-1} r_i^{\beta_i} w_i^{1-\beta_i} \quad (4)$$

Marginal costs c_i are assumed constant and equal to the cost per unit of output. Marginal costs are derived from the exogenous parameters technology A_i , real wage w_i , real return rate on investment r_i and the Cobb-Douglas function exponents β_i and $1-\beta_i$. If the returns on investment and the exponents are similar in the East and West, a

⁴ The Bazaar Economy hypothesis has been rejected by other authors, for instance Horn and Behncke (2004).

cost advantage can be achieved via a more efficient technology or a lower wage level. X producers maximise their profits and have oligopolistic (monopolistic) power reflected in the price for X depending on the quantity of X , where $X = X_W + X_E$:

$$\max \Pi_i = p_i(X) \cdot X_i - c_i \cdot X_i \quad (5)$$

Without any market power of X producers prices would be equal to marginal costs. In general a productivity gap between the East and the West exists, and hence, under perfect competition intermediate good production would take place in the area with lower marginal costs only.

In the case of Bertrand competition (price competition) intermediate good manufacturing occurs only in the cheaper region, too. Now the cheaper producer can increase the price for X and reduce the production quantity of X . But if he increases the price for X more than to the marginal costs of the rival X producer, he will lose all the demand for his product.

Cournot competition (competition in quantities) is the interesting case referring to the current tendencies in the European automotive industry as well as analytically.

4 Cournot Competition

Production of automotive components, for example seats or dashboards containing numerous distinct parts, requires low-skilled manufacturing. Nevertheless, specific knowledge and a sufficient firm size are necessary to produce automobile parts in large amounts according to the final producer's needs and fulfilling the quality requirements. The intermediate good and the final good company agree to a long term contract, so that other suppliers can hardly enter the market. For that reason it is plausible to suppose a small number of companies being able to provide the specific components needed in final production, in this model in particular two suppliers. Under the assumption of pure Cournot competition the Eastern and Western company offer the intermediate good X at the same price $p_X = p_W = p_E$. The firms optimise their supply of X taking into account the rival's reaction and the demand function for X given by the Western final good Y producer. The conditional factor demand functions can be derived from (2) in the standard way:

$$X(Y, w_H, p_x) = \left(\frac{\alpha p_x}{(1-\alpha)w_H} \right)^{-\alpha} Y \quad (6)$$

Total demand for X falls with the price p_x and increases with w_H , the price of the skilled labour intensive good H . Similarly the input quantity of H is expressed:

$$H(Y, w_H, p_x) = \left(\frac{\alpha p_x}{(1-\alpha)w_H} \right)^{1-\alpha} Y \quad (7)$$

A change in the exogenous quantity Y leads to proportional shifts of the input factors X and H . (The absolute value of Y is not important when dealing with relative values between the East and the West, since Y drops out). Solving (6) for p_x yields the inverse factor demand function for X :

$$p_x = \frac{1-\alpha}{\alpha} w_H X^{-\frac{1}{\alpha}} Y^{\frac{1}{\alpha}} \quad (8)$$

Obviously, the intermediate good suppliers face a non-linear and downward-sloping inverse factor demand function with respect to the total production of X . Since $\alpha < 1$, expanding the supply of total X leads to a more than proportional fall in the price p_x . Hence, a monopolist would choose the output as small as possible, but in the oligopoly the situation is different. For every given positive quantity of one supplier there is an optimal output of the rival, which results in an equilibrium with positive quantities. In market equilibrium supply equals demand for X at the price p_x , so that we can insert (8) into (5) recalling that the final producer is a price taker:

$$\max \Pi_i = \frac{1-\alpha}{\alpha} w_H X^{-\frac{1}{\alpha}} Y^{\frac{1}{\alpha}} \cdot X_i - c_i \cdot X_i \quad (9)$$

The oligopolists maximise their profits Π_i by choosing their production quantities X_i and taking into account the total amount X that includes their own and their rival's quantity. This results in the following first order conditions for profit maximisation representing the oligopolists' reaction functions with $i = [W; E]$:

$$\frac{d\Pi_i}{dX_i} = \frac{1-\alpha}{\alpha} w_H Y \alpha \left(-\frac{1}{\alpha} X^{-\frac{1}{\alpha}-1} \cdot X_i + X^{-\frac{1}{\alpha}} \right) - c_i = 0 \quad (10)$$

The appendix⁵ (a) shows that the second order condition for a profit maximum is fulfilled. Dividing the first order conditions for the East and the West yields the intermediate good output ratio (appendix b):

$$V_{WE} \equiv \frac{X_W}{X_E} = \frac{(1-\alpha)c_{WE} + \alpha}{(1-\alpha) + \alpha c_{WE}} \quad (11)$$

V_{WE} , a key variable in this model, is the ratio of production in the West relative to the East. c_{WE} means c_W / c_E , that is marginal production costs in the West divided by marginal costs in the East, in other words the relative Western marginal costs.

5 Allocation of Intermediate Good Production

This section examines how the allocation of intermediate good production to Eastern and Western Europe reacts to changes in the production cost differential between those regions. For this purpose the elasticity of relative interregional production is defined. The first derivative of expression (11) with respect to marginal costs in the West relative to the East shows the reaction of relative X production V_{WE} to changes in relative marginal costs c_{WE} (more detailed in the appendix b).

$$\frac{dV_{WE}}{dc_{WE}} = \frac{1-2\alpha}{[(1-\alpha) + \alpha c_{WE}]^2} < 0 \quad (12)$$

α is larger than 0.5 per plausible assumption, and hence the term above is negative. That means, increasing marginal production costs in the West relative to the East lower the relative Western production V_{WE} as expected. Facing higher production costs in the West, the oligopolist reduces the profit maximizing output X_W , so that the relative Western X output falls. Now the following elasticity ε can be derived from

⁵ The mathematical appendix is available upon request, http://www.ifw-kiel.de/research/the-environment-and-natural-resources/view?set_language=en.

(11) and (12) in order to analyse the sensitivity of relative X production to changes in relative production costs c_{WE} :

$$\varepsilon = \frac{dV_{WE}}{dc_{WE}} \cdot \frac{c_{WE}}{V_{WE}} = \frac{1-2\alpha}{[(1-\alpha) + \alpha c_{WE}]^2} \cdot \frac{c_{WE}}{\frac{(1-\alpha)c_{WE} + \alpha}{(1-\alpha) + \alpha c_{WE}}} < 0 \quad (13)$$

This elasticity of relative interregional production is a measure for the sensitivity of offshore outsourcing to interregional production cost changes. When it is high, cost deviations at home lead to a large relative production shift to abroad. Term (13) and the graphical representation in Figure 1 show that the elasticity of relative intermediate good production is a function of relative production costs. (Note that the graph is symmetric, that is the branch left of the minimum is the inverse of the branch on the right hand side of the minimum, so that the results are independent of looking at c_W / c_E or at c_E / c_W). The elasticity curve is steep near its extremum, while it is flatter, when relative production costs deviate more from 1. We now analyse the first derivative of the denominator of equation (13) in order to find an extremum of ε :

$$\frac{d}{dc_{WE}} \left[(1-\alpha)^2 + \alpha^2 + \alpha(1-\alpha) \frac{1}{c_{WE}} + \alpha(1-\alpha)c_{WE} \right] = 0 \Rightarrow c_{WE} = 1 \quad (14)$$

It can be shown that ε has a minimum lower than zero for $c_{WE} = 1$ (appendix b). For c_{WE} towards infinity or towards zero the elasticity goes towards zero (Figure 1).

With the definition of V_{WE} according to (11) in the middle bracket and some algebra (see appendix, c) X_E and X_W can be expressed as:⁶

$$X_E = \left(\alpha' \frac{w_H}{c_W + c_E} \right)^\alpha \frac{1}{1+V_{WE}} Y; \quad X_W = \left(\alpha' \frac{w_H}{c_W + c_E} \right)^\alpha \frac{1}{1+\frac{1}{V_{WE}}} Y; \quad (15)$$

$$\text{where } \alpha' = \frac{(1-\alpha) \left(2 - \frac{1}{\alpha} \right)}{\alpha}$$

We write α' for simplicity. $1 / (1 + V_{WE})$ is the Eastern European share of X production, and $1 / (1 + 1/V_{WE})$ is the Western share of total X production. The Eastern supply of X increases proportionally with the exogenous final output Y . The

⁶ The errors in the published article have been corrected.

higher the sum of marginal costs in the West and the East compared to the price for high-skilled labour w_H , the lower the input of X_E and also the input of X_W in final production.

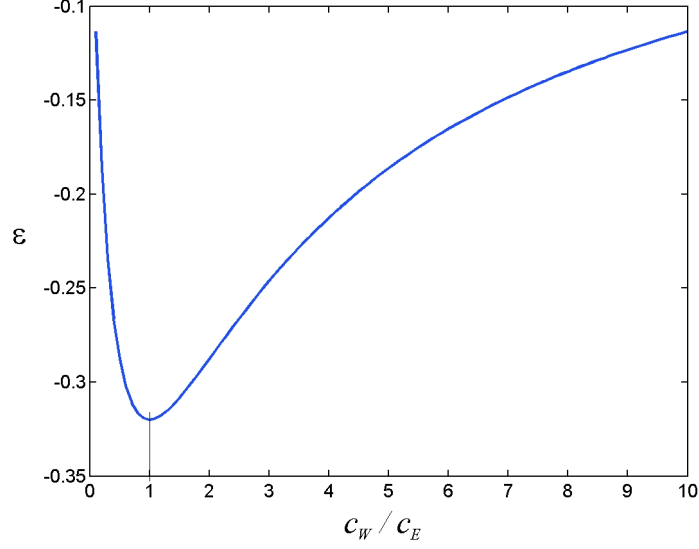


Figure 1: Elasticity ε of intermediate good production in the West relative to the East (V_{WE}) with respect to relative production cost changes (c_{WE}), example with halflogarithmic scale, $\alpha = 0.66$

Equation (10) represents a Nash equilibrium, because both oligopolists maximise their profits taking the rival's behaviour into account, and the firm that deviates from this optimal production will suffer a lower profit. Adding the Eastern and the Western output derived from (10) leads to total supply of X :

$$X = X_W + X_E = \left(\alpha' \frac{w_H}{c_W + c_E} \right)^\alpha Y \quad (16)$$

The equilibrium price p_X which is identical for X_W and X_E per assumption, becomes obvious by comparison of the above term with expression (6). It follows:

$$p_X = \frac{c_W + c_E}{2 - \frac{1}{\alpha}} \quad (17)$$

The equilibrium price of X is a linear function of the sum of the marginal costs in the regions. The input of X as a production factor in final production decreases when p_X

risers. Consequently, higher costs in one region not only reduce the output in this region, but also total supply of the intermediate good X by increasing p_X . Hence it is not immediately clear, if higher production costs in the East increases or decreases absolute X production in the West. This depends on whether the production shift from East to West is lower or higher than the negative effect of the total X production decline. The magnitude of p_X movements due to marginal cost changes is determined by the coefficient $1 / (2 - 1 / \alpha)$. For α approaching 1, p_X is approximately the sum of marginal costs. This would imply a very high exponent in the Cobb-Douglas function for H and a very low exponent for the input X . The magnitude becomes higher and higher, when α falls towards 0.5, keeping in mind, that 0.5 is a lower bound for α in this model.

The profits of the intermediate good suppliers can easily be expressed with the help of the price p_X :

$$\Pi_i = p_x \cdot X_i - c_i \cdot X_i \quad (18)$$

As the simulation will show, profits decrease with increasing costs as expected. But not only the profits of the producer facing rising costs do fall, the other supplier's profits also slightly fall together with a small output reduction.

6 Allocation of Production Factors

This section describes the profit maximizing allocation of production factors depending on production costs. We concentrate on labour inputs, while absolute and relative capital inputs can be derived in an analogue way. The analytic relationships are then used to carry out simulations in section 7.

At first we look at the input of high-skilled labour intensive input H by plugging p_X into (7):

$$H(Y, w_H, c_W, c_E) = \left(\frac{1}{\alpha'} \frac{c_W + c_E}{w_H} \right)^{1-\alpha} Y \quad (19)$$

It follows for the ratio of H to X :

$$\frac{H}{X} = \frac{\frac{p_X}{1-\alpha}}{\frac{w_H}{\alpha}} = \frac{1}{\alpha'} \frac{c_W + c_E}{w_H} = \frac{c_W + c_E}{1-\alpha} \frac{\alpha}{w_H} \frac{1}{2 - \frac{1}{\alpha}} \quad (20)$$

This cost minimizing ratio of high-skilled labour intensive input H to the low-skilled manufactured input X is the standard microeconomic outcome, but with p_X expressed as shown before. A rising sum of marginal costs in X manufacturing shifts the input intensity in final production from X to H . In contrast to the standard microeconomic relationship the impact of α is unclear, as α also affects p_X via the factor $1 / (2 - 1 / \alpha)$.

At second we write the standard conditional factor demands for low-skilled labour L_i derived from the production function (3) with $i = [W; E]$ in order to produce the amounts X_i at minimal costs (Varian 2001). Note that capital and labour are specific for each region, that is immobile between the regions, so that there is not factor price equalisation, which implies a short- or medium-run point of view.

$$L_i(X_i, r_i, w_i) = \frac{1}{A_i} \left(\frac{(1-\beta_i)r_i}{\beta_i w_i} \right)^{\beta_i} X_i \quad (21)$$

The ratio of labour inputs is expressed using the definition of V_{WE} according to (11):

$$\frac{L_W}{L_E} = V_{WE} \frac{A_E}{A_W} \frac{\left(\frac{\beta_E w_E}{(1-\beta_E)r_E} \right)^{\beta_E}}{\left(\frac{\beta_W w_W}{(1-\beta_W)r_W} \right)^{\beta_W}}, \quad (22)$$

$$\text{where } V_{WE} = f \left(\frac{-}{c_E} \right) = f \left(\frac{-}{A_W}; \frac{-}{w_E}; \frac{-}{r_E} \right)$$

The minus signs indicate that V_{WE} is a falling function of A_E / A_W , w_W / w_E and r_W / r_E . The labour input in the West compared to the East is therefore determined by the relative productivity, the relative Western return rate on investment and the relative Western wage level. Within a region it depends on the price of capital relative to the price of labour as usual. When $\beta = \beta_W = \beta_E$, the relation simplifies to:

$$\frac{L_W}{L_E} = V_{WE} \frac{A_E}{A_W} \left(\frac{\frac{w_E}{r_E}}{\frac{w_W}{r_W}} \right)^\beta = V_{WE} \frac{A_E}{A_W} \left(\frac{w_E}{w_W} \frac{r_W}{r_E} \right)^\beta \quad (23)$$

A higher relative Western wage clearly lowers the labour input in the West in comparison to the East (directly in 22 and indirectly via V_{WE}), but the impacts of changes in r_W / r_E and A_W / A_E are per se not clear. On the one hand, a higher Western return rate on capital shifts production towards more labour input in the West; on the other hand, a higher return rate increases overall production costs, which drives production and thus labour and capital inputs from the West to the East. The result of a productivity improvement in the West is ambiguous, too: It lowers production costs extending Western X production, but at the same time a given output Y can be produced with less labour and capital inputs.

To conclude, the analysis has shown that the wage for high-skilled workers w_H compared to the sum of marginal production costs in the East and West determines the ratio of H to X in final good production, while the marginal cost ratio c_W / c_E determines the distribution of X production to the West and to the East.

7 Transport Costs

Manufacturing the intermediate good X in the Eastern region and importing it for final production into the Western region creates transportation costs for X . Referring to the classic approach by Samuelson (1954), transport costs C_T are represented by a quantity melting like ice. For that purpose an additional factor $(1 - C_T)$ is introduced in (3) and for simplicity combined with the productivity coefficient A_E :

$$X_E = A'_E K_E^{\beta_E} L_E^{1-\beta_E}; \quad \text{where} \quad A'_E = (1 - C_T) A_E; \quad 0 \leq C_T < 1 \quad (24)$$

It is immediately obvious that transport costs lower the Eastern productivity or in other words increase Eastern marginal production costs and hence shift production from the East to the West. Transport costs will be included in the simulations in the next section.

8 Graphical Representations

The following graphs visualise the effects of a change in relative production costs in form of a rising Western wage level. The simulations are based on the equations derived in sections 5, 6 and 7 and the analogue expressions for capital inputs.

Figure 2 illustrates the change in absolute amounts of X production in the East and the West and the corresponding profits Π_W and Π_E (according to equations 15 and 18), when the Western wage rises relative to the Eastern wage, while the Eastern wage is fixed.⁷

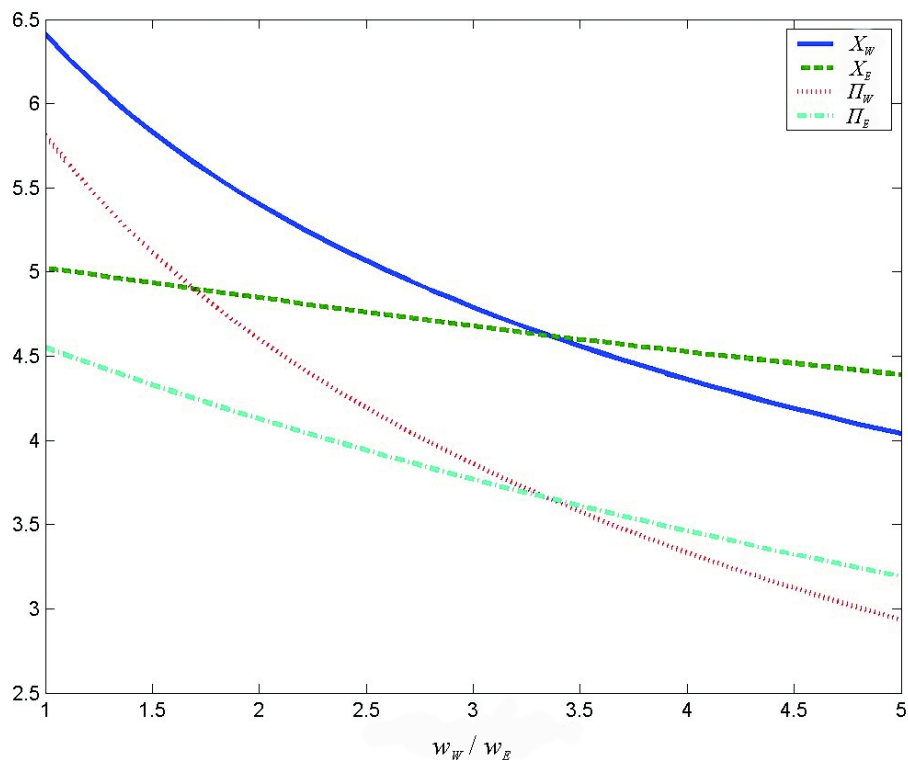


Figure 2: Absolute production allocation depending on an increasing western wage

The following assumptions on parameter values have been made for the functions of the model:⁸ For any given relative wage the graphs show the inputs necessary to produce one unit of final output Y . The Cobb-Douglas exponent for the high-skilled labour intensive input α equals 0.66 and that one for the X input 0.34. The price of H is set 15 times higher than the payment for Eastern low-skilled workers ($w_E = 1$).

⁷ The profits are divided by 10 to scale them down in the figure.

⁸ For a comparison of relative wages, productivity and labour costs between Germany and Austria and Eastern European regions see Marin (2004).

Accordingly, there is a much higher income share for high-skilled workers such as managers, engineers and designers creating H than for low-skilled workers manufacturing X in the East. In the graph the Western wage (w_W) rises exogenously from 1 up to 5 relative to the Eastern wage. For example the labour cost ratio between Germany and the Czech Republic amounts to about 5 and might decrease in the future. Returns on capital in the East and West are set equal to one, since the graph focuses on wage differentials. The exponents β_i related to capital inputs in intermediate good production are assumed to be 0.34 in both regions. The exponents $1-\beta_i$ for labour inputs consequently lead to higher income shares for low skilled-workers than for capital owners in every region. The Western X producer has an advantage in total factor productivity, for simplicity A_E is set to 1 and A_W to 2. The intuition is access to better production technologies in the West. Finally transport costs C_T are included amounting to 10% of the transferred good X causing a further cost disadvantage for the East.

Consequently, for equal wages ($w_W / w_E = 1$, left hand side of Figures 2 and 4) the manufactured quantity of X in the West is higher than in the East because of the higher Western productivity level and the transport costs of importing X_E from the East to the West. As expected, rising the Western wage while holding the Eastern wage constant reduces Western X production strongly. Surprisingly, at the same time Eastern X production slightly decreases resulting in an even larger decrease in total X supply. The reason is that higher Western production costs increase p_X (the common price for Eastern and Western good X) proportionally as shown in equation (17). When the price for the intermediate good X increases, the final good producer replaces X by H leading to a (substantial) decline in the demand for X . Hence, total X , which is the sum of X_W and X_E , also falls. Corresponding to the quantities Western profits decrease strongly with rising labour costs, and Eastern profits fall slightly.

Figure 3 points the mechanisms on the X supply side out. The curves are the graphical representations of the reaction functions in (15) of the Western and the Eastern firm, each representing profit maximizing output as a function of the rival's output.

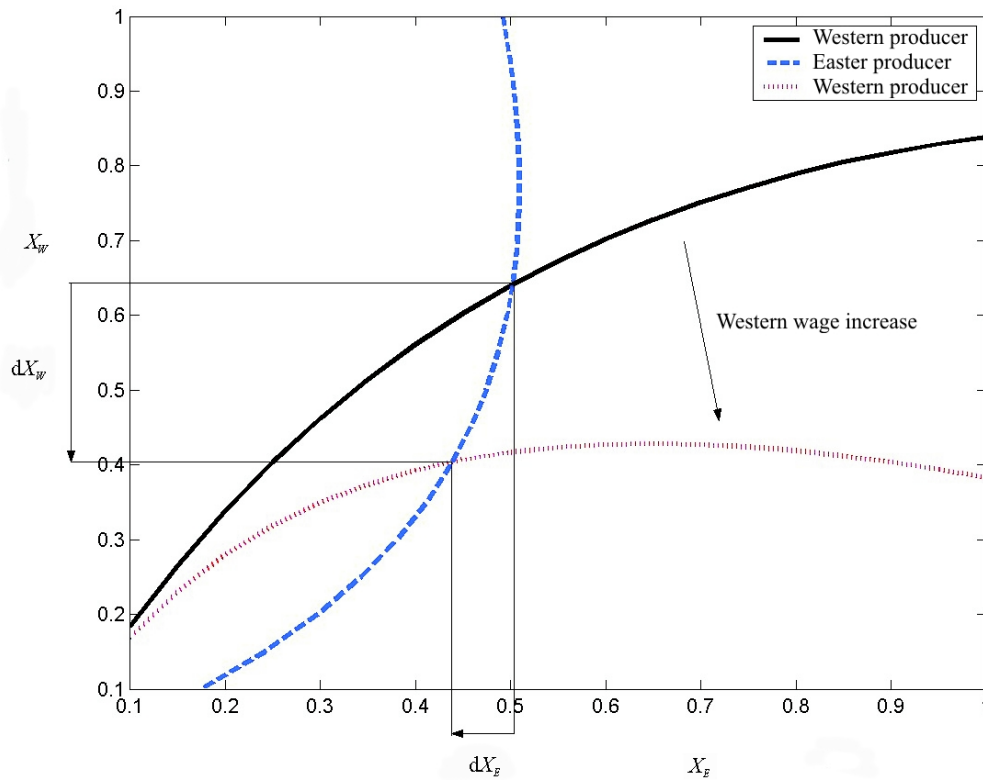


Figure 3: Profit maximizing reaction functions

The parameter values are set as before. The curves are again asymmetric because the Eastern oligopolist faces transport costs and an inferior technology. Obviously, the reaction curves are upward sloping, so that one firm reacts to a higher quantity of the other firm with another output expansion. In certain regions of the curves and for certain parameter values it is also possible to find downward sloping or approximately vertical or horizontal reaction curves. Then the firm reacts to the rival's output expansion with an output reduction or is not affected at all. But the typical case in this model framework is that of upward sloping reaction curves causing the important result of an output reduction in both regions for a cost increase in one region. This means, the intermediate good production factor is replaced by the high-tech production factor. Is a zero-zero solution without intermediate good production feasible? The quantities X_W and X_E are endogenous and will never become jointly zero, nor will one quantity become zero. This means, the model cannot represent the non-production case and the monopoly case. Even when a very large inter-regional cost differential exists, there will be a marginally small rest of production in the high-cost region.

As shown in Figure 3, a Western wage raised from 1 to 5 relative to the Eastern wage shifts the Western reaction curve downwards lowering Western output strongly (dX_W) and decreasing Eastern output slightly (dX_E).

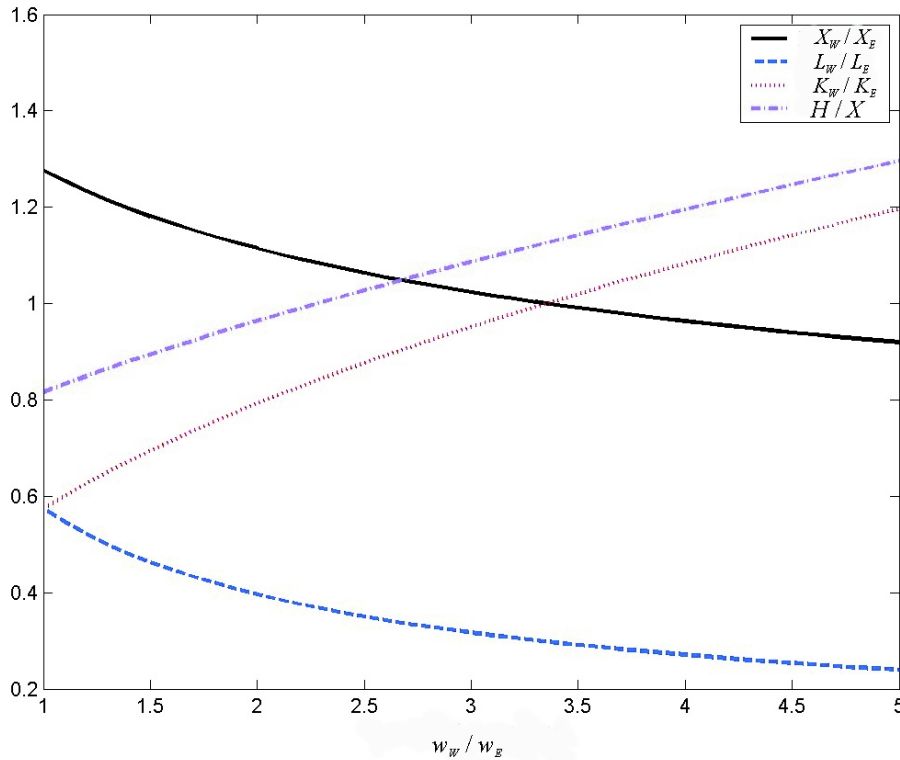


Figure 4: Relative factor allocation depending on an increasing western wage

Figure 4 plots the relative factor distribution, that is Western in relation to Eastern X production, Western compared to Eastern low-skilled labour and capital inputs as well as the ratio of the Western high-skilled labour intensive factor H to low-skilled labour intensive X being produced in the West and the East (equations 11, 21 and similarly for capital inputs). Again the Western wage is exogenously increased relative to the fixed Eastern wage.

All parameter values are the same as before and the results are in accordance with those of Figure 2. For the case of identical wages (left hand side) the output of X in the West is higher than in the East ($X_{WE} > 1$) because of the Western technological advantage and the Eastern transport costs disadvantage. Nevertheless, low-skilled labour input is lower in the West than in the East, and capital input is lower in the West, too. This follows from the fixed final good Y output limiting the demand for factors H and X and from the Western productivity advantage reducing the required

input quantities to produce a certain output quantity.

Obviously all curves in figures 2 and 4 have a sharper (rising or falling) slope when labour costs are similar (left hand side) in accordance with the higher elasticity of relative production for equal marginal costs derived in chapter 5 (equation 13). This is caused by the strictly convex inverse factor demand function (8) which is steeply falling for low prices in combination with the reaction functions (15).⁹

9 Interpretation and Caveats

The model analysis has the following implications: It reveals that the elasticity of interregional outsourcing activities as a reaction to interregional cost changes depends on the original gap in production costs between the regions. When intermediate good production occurs in two regions with similar production costs, any cost change in one region will have strong re-allocation effects. Hence, for example a higher wage agreement for the Western intermediate good firm or a higher wage level after bargaining with unions lowers the Western intermediate good production relative to the Eastern one a lot in terms of relative changes. In the same way higher additional labour costs (taxes and insurance) have a strong negative influence on Western production and labour input. Given a situation of a large East-West production cost differential and completed offshore outsourcing activities, additional cost changes will have little effect on outsourcing, thus contradicting outsourcing fears.

Moreover, higher marginal costs in one region not only reduce intermediate good output in this region, but also in the other region due to a rising common intermediate good price, which reduces demand for the intermediate good. Thus, the cost increase reduces intermediate good production in both regions. Of course the output decline is higher in the region, where the cost increase occurs. In order to keep final output constant low-skilled labour intensive intermediate good input is substituted by the high-skilled labour (and advanced technology rich) input located in Western Europe. Consequently, any rise in intermediate good production costs benefits Western high-skilled workers. The outcome for low-skilled workers in the region with rising costs depends on whether the cost increase includes higher wages. In case of higher wages the entire group of low-skilled workers in that region can still

⁹ Since a Cobb-Douglas function with constant returns to scale implies a decreasing marginal product of labour, the inverse factor demand function has a convex falling shape.

lose because the production decline due to higher costs eliminates jobs. If the wage increase overcompensates the reduced labour input, the workers can all be better off after redistribution. However, this is unlikely the case. We saw that a cost increase in one region also slightly reduces intermediate production in the other region. Therefore, *ceteris paribus* capital owners and workers involved in intermediate production in the other region also lose to some extent. We conclude that there is no direct competition between low-skilled workers in the East versus the West, but mainly between Western high-skilled on the one hand and both Eastern and Western low-skilled workers on the other hand in terms of absolute input quantities. Of course, referring to relative quantities, any cost increase in one region lowers the relative production share of this region compared with the other region. According to the graphical interpretation for a specific parameterisation, an advantage via a higher productivity leads to a higher intermediate good output with relatively lower factor input quantities.

When applying and interpreting the model several caveats should be considered: The paper is a typical partial equilibrium analysis. It abstracts from trade and terms of trade effects. Therefore, it is not possible to carry out a general welfare analysis.

The model concentrates on the supply side, assuming a constant demand for Y . The analysed changes in factor prices or production technologies do not affect the price for Y , either. But since the development of consumer tastes and technological change during the allocation process are unknown, it seems reasonable to set these variables exogenous and constant.

An oligopoly in intermediate good production is a sensible assumption for certain sectors of the economy, but of course not for all sectors. The automobile industry has been chosen as an example. However, the oligopoly assumption would also be sensible in final good production, which has been neglected here in order to keep the model analytically tractable. Furthermore, a typical automobile producer has to some extent market power concerning the purchase of parts and components. This model is in favour of the intermediate good suppliers concerning market power.

Furthermore, the findings are based on convex decreasing inverse factor demand functions derived from Cobb-Douglas technologies. Thus, the intermediate good producers tend to keep their output low for the purpose of holding the price high. The results probably also hold when assuming CES production functions. However, the

substitution effect between high-skilled and low-skilled labour will be stronger, when the elasticity of substitution of the CES function is higher. On the other hand, a higher elasticity of substitution between labour and capital in intermediate production will probably make the outsourcing behaviour less sensitive to labour cost changes, because capital can more easily substitute for labour within intermediate in one region.

The exponent of the high-skilled labour intensive input in the Cobb-Douglas function is assumed to be larger than one half, so that the exponent of the low-skilled labour intensive input is smaller than one half. This necessary assumption seems sensible, since it implies a higher income share for the combination of high-skilled labour and high technology capital found in the Western region than for the low-skilled labour intensive input that is outsourced. It indicates in a way that production is mainly located in the Western region.

When a very large inter-regional cost differential exists, there will still be a marginally small rest of production in the high cost region. In that sense, the Cournot competition assumption is not realistic for very large cost differentials. On the other hand, a large cost differential points to a big difference in per capita incomes and the levels of development. Hence, poor infrastructure, unsecured property rights, corruption and other disadvantages and risks dominate the outsourcing decision of investors rather than pure production costs. These factors prevent higher outsourcing into the low cost country even when a large cost differential exists. Moreover, the results show the strongest elasticity of outsourcing activities to production cost changes when the interregional cost difference is smallest, that is the countries are similar. In reality there is a sluggishness of production movements due to the home bias and costs of planning and organizing the outsourcing activity. This sluggishness opposes the offshore outsourcing incentive, and the outsourcing costs create a threshold, that is a minimum cost differential, which is necessary to cause any offshore outsourcing.

Do the results of the simulations hold for all parameter values? Actually, it is possible to find certain parameter values that yield the classic outcome: One oligopolist reduces output due to higher production costs and the other firm reacts with an output expansion. Another possibility is a reaction function that is locally vertical or horizontal. Then the first firm can change output without affecting the second firm. Nevertheless, these are special cases, while the normal behaviour for

most parameter values within this model framework is as explained before.

Finally, a precondition for the analysis is a situation of completed adjustment of factor allocation according to efficiency or profit maximisation, respectively. When applying the model to the current situation of the European Union we need to take into account, that dynamic adjustment processes are still in progress. Consequently, adjustment processes towards the equilibrium offset the outcomes resulting from the model. The model does not assume factor price equalisation. This means, it has a medium-run point of view, or there are persistent barriers to trade or factor movements.

10 Conclusion

This paper deals with the question, how and how sensitively outsourcing from a high-technology area with high production costs like Western Europe to a low-technology and low-cost area like Eastern Europe reacts to changes in production costs. It refers to current tendencies in the (European) automotive industry: Concentrations of firms in final production as well as in production of parts and components, increasing market power of the suppliers, but at the same time price or cost pressure, and offshore outsourcing to Eastern European countries with lower wage levels.

The model presented here is a two stage model based on Cobb-Douglas technologies and Cournot competition in intermediate goods as perfect substitutes. Final good production takes place in Western Europe, while the intermediate activity can be located in Western Europe (outsourcing) or Eastern Europe (offshore outsourcing).

It is shown how relative production costs, determined by technology levels, wages and returns on investment, influence the allocation of production and of labour input between the two regions. The first key result is, that intermediate good production falls in region A relative to B, when production costs in A rise relative to B. But intermediate good output in B does also slightly fall in absolute terms as a result of the Cournot competition. Graphical simulations illustrate the findings. Thus, there is a substitution effect between high-skilled and low-skilled labour rather than between low-skilled labour in the East and in the West.

The second key result is that the sensitivity of outsourcing behaviour to production cost changes is higher the smaller the cost difference between the regions

is. Consequently the effect of changes in wage agreements or labour taxes depends on the original cost gap. Given a situation of a large East-West production cost differential and completed outsourcing activities, additional cost changes will have little effect on outsourcing behaviour. This contradicts the fear of production rapidly moving towards East once Western wages or labour taxes change slightly. Moreover, the simulation example shows that a superior technology in one region leads to a higher relative output and lower relative factor inputs in that region compared to the other region.

Currently labour costs differ strongly between Eastern European countries. According to this simplified model and ignoring other factors driving outsourcing decisions, the sensitivity of Western outsourcing activities to production cost changes in countries like Czech Republic and Hungary is higher than in Romania and Bulgaria, since the wage levels of the former countries are closer to the Western wage level than those of the latter countries. When the Eastern production cost level converges to the Western level, the outsourcing sensitivity will increase. However, the analytical model does not capture outsourcing determinants like infrastructure and the legal framework predicting the most offshore outsourcing into countries with the lowest labour costs.

The model serves as an easily tractable analytical tool and can be implemented in a more detailed general equilibrium model. Particularly, the final good demand side and international trade can be implemented. Income effects for high- and low-skilled workers and capital owners in the two regions can then be investigated in a more detailed way, and a general welfare analysis can be done.

References

Baldwin, R. and F. Robert-Nicoud (2007). Offshoring: General Equilibrium Effects on Wages, Production and Trade, NBER Working Paper No. 12991.

Bandyopadhyay, S. and H. J. Wall (2005a). Immigration and Outsourcing: A General Equilibrium Analysis, IZA Discussion Paper No. 1694.

Bandyopadhyay, S. and H. J. Wall (2005b). Oligopoly and Outsourcing, IZA Discussion Paper No. 1892.

Bartel, A. P., Lach S. and N. Sicherman (2005). Outsourcing and Technological Change, CEPR Discussion Paper No. 5082.

- Becker, S. O., K. Ekholm, R. Jäckle and M. A. Münder (2005). Location Choice and Employment Decisions: A Comparison of German and Swedish Multinationals, *Review of World Economics* 141(4), 693-731.
- Bhagwati, J., A. Panagariya and T. N. Srinivasan (2004). The Muddles over Outsourcing. *Journal of Economic Perspectives* 18 (4), 93-114.
- Boeri, T. and H. Brücker (2000). The Impact of Eastern Enlargement on Employment and labour in the EU Member States, Final Report, European Integration Consortium, DIW/CEPR/FIEF/IAS/IGIER.
- Braconier, H. and K. Ekholm (2001). Foreign Direct Investment in Central and Eastern Europe, CEPR Discussion Paper No. 3052.
- Dunning, J. (1981). *International Production and the Multinational Enterprise*, London, George Allen and Unwin.
- Egger, H. and P. Egger (2004a). On the Relationship between International Outsourcing and Price–Cost Margins in European Industries, *Review of Industrial Organization* 25, 45–69.
- Egger, H. and P. Egger (2007). Outsourcing and Trade in a Spatial World, *Journal of Urban Economics* 62(3), 441-470.
- Feenstra, R. C. and G. H. Hanson (1996a). Foreign Investment, Outsourcing, and Relative Wages, in: Feenstra R. C., G. M. Grossman and D. A. Irwin (eds.), *The Political Economy of Trade Policy – papers in honor of Jagdish Bhagwati*, MIT Press.
- Feenstra, R. C. and G. H. Hanson (1996b). Globalization, Outsourcing, and Wage Inequality, *American Economic Review* 86, 240-245.
- Grossman, G. M. and E. Helpman (2002). Integration versus Outsourcing in Industry Equilibrium, *Quarterly Journal of Economics* 117(468), 85-120.
- Grossman G. M. and E. Helpman (2003). Outsourcing versus FDI in Industry Equilibrium, *Journal of the European Economic Association* 1(2), 317-327.
- Grossman, G. M. and E. Helpman (2005). Outsourcing in a Global Economy, *Review of Economic Studies* 72(1), 135-159.
- Grossman, G. M. and E. Rossi-Hansberg (2006a). Trading Tasks: A Simple Theory of Offshoring, NBER Working Paper No. 12721.
- Grossman, G. M. and E. Rossi-Hansberg (2006b). The Rise of Offshoring: It's Not Wine for Cloth Anymore, Federal Reserve Bank of Kansas City, Economic Symposium Proceedings, *The New Economic Geography: Effects and Policy Implications*.
- Horn, G. and S. Behncke (2004). Deutschland ist keine Basarökonomie, *Wochenbericht des DIW Berlin* No. 40/04.
- Kirkegaard, J. F. (2005). Outsourcing and Offshoring: Pushing the European Model Over the Hill, Rather Than Off the Cliff!, *Institute for International Economics*

Working Paper No. 05-1.

Kohler, W. (2004). International Outsourcing and Factor Prices with Multistage Production, *The Economic Journal* 114, C166–C185.

Konings, J. and A. Murphy (2001). Do Multinational Enterprises Substitute Parent Jobs for Foreign Ones? Evidence from European Firm-Level Data, CEPR Discussion Paper No. 2972.

Marin, D., A. Lorentowicz and A. Raubold (2002). Ownership, Capital or Outsourcing: What Drives German Investment to Eastern Europe?, Munich Economics Discussion Paper No. 02-03.

Marin, D. (2004). A Nation of Poets and Thinkers – Less so with Eastern Enlargement? Austria and Germany, CEPR Discussion Paper No. 4358.

Marin, D. (2006). A New International Division of Labour in Europe: Outsourcing and Offshoring to Eastern Europe, *Journal of the European Economic Association* 4(2-3), 612–622.

Markusen, J. R. and K. E. Maskus (2001). Multinational Firms: Reconciling Theory and Evidence, in: Blomstrom M. and L. S. Goldberg, *Topics in Empirical International Economics: A Festschrift in Honor of Robert E. Lipsey*, University of Chicago Press.

Markusen, J. R. (2005). Modeling the Offshoring of White-Collar Services: From Comparative Advantage to the New Theories of Trade and FDI, NBER Working Paper No. 11827.

Meissner, H. R. and U. Jürgens (2007). Zur Lage der deutschen Automobil-Zulieferindustrie im Jahr 2007, Eine Studie im Auftrag der Bundestagsabgeordneten Sabine Zimmermann, Fraktion Die Linke.

Mitra, D. and P. Ranjan (2005), Y2K and Offshoring: The Role of External Economies and Firm Heterogeneity, NBER Working Paper No. W11718.

Munch, J. R. and J. R. Skaksen (2005). Specialization, Outsourcing and Wages, IZA Discussion Paper No. 1907.

Samuelson, P. A. (1954). The Transfer Problem and Transport Costs: Analysis of Effects of Trade Impediments, *Economic Journal* 64(254), 264-289.

Senses, M. Z. (2006). The effect of Outsourcing on the Elasticity of Labour Demand CES No. 06-07.

Sinn, H. W. (2005). Basar-Ökonomie Deutschland – Exportweltmeister oder Schlusslicht?, ifo Schnelldienst 6.

Varian, H. R. (2001). *Mikroökonomie*, 3. Auflage, München, Wien, R. Oldenburg Verlag.

Wang, B. (2006). Vertical Integration versus Outsourcing in Industry Equilibrium, CCER Peking University Working Paper No. E2006006.

Chapter 1:

The Labor Market Effects of Outsourcing Parts and Components

Part 2: Adverse Cournot Competition *

Michael Hübler ⁺

October 13, 2009

Abstract

This paper contributes to Hübler (2008) who analyses a partial equilibrium model of outsourcing with Cournot competition in intermediate good production. Final production is located in Western Europe, whereas the intermediate good can be manufactured by a Western (outsourcing) or Eastern European supplier (offshore outsourcing). The paper asks the question how changes in production costs, in particular wages, affect output and thus labor input in the two regions. The paper proves analytically that under certain conditions higher production costs in one region reduce intermediate good production in both regions.

JEL Classifications: D24, D43, F20, J31

Keywords: Offshoring, outsourcing, Cournot competition,
intermediate goods, high-skilled, low-skilled

* I thank Johannes Bröcker for very helpful comments.

⁺ Kiel Institute for the World Economy, Düsternbrooker Weg 120, 24100 Kiel, Germany, E-mail: michael.huebler@ifw-kiel.de, Tel: +49-431-8814-401.

1 Introduction

A main result in Hübler (2008) is the possibility of “adverse Cournot competition”. This means, one producer decreases output as a reaction to an output decrease of the other producer. Thus, if the first producer faces a cost increase, he will decrease his output, and the second producer will also do so, however to a smaller extent, so that the second producer will produce more relative to the first producer. While Hübler (2008) derives this result from numerical simulations, this paper derives the necessary conditions for this outcome strictly analytically in section 4. Section 2 repeats the model setup described by Hübler (2008). Section 3 provides the proof that the second order condition for a profit maximum in Hübler (2008) is indeed fulfilled. Section 5 contains corrections of Hübler (2008). Section 6 concludes.

2 Model setup

The partial equilibrium model set up in Hübler (2008)¹ consists of final good Y production, located in the Western European region, and intermediate good X production located in the Western and in the Eastern European region.

$$Y = H^\alpha (X_W + X_E)^{1-\alpha}; \quad 0.5 < \alpha < 1 \quad (1)$$

The final good producer takes demand as given by the market. (The final good market form is not of importance in this case.) H is a high-skilled labor and high-technology intensive input, which is available in the Western area only and cannot be outsourced.

The process of intermediate good X production includes all activities requiring low-skilled labor like manual work and usual capital. X can be manufactured in Western (X_W) or Eastern Europe (X_E). Low-skilled labor is supplied in both regions. Intermediate good production can be moved to the East when production costs or wages are cheaper in the East, afterwards the manufactured intermediates are imported to the West. X_W and X_E are homogeneous goods and perfect substitutes so that they can be summed up to $X = X_W + X_E$. The assumption $\alpha > 0.5$ implies a higher income share for the high-skilled labor and high-technology intensive input located in

¹ This section is partly a citation of Hübler (2008), section 3.

the West (under perfect competition). The Cobb-Douglas function implies the possibility of replacing part of input X , produced with a high amount of low-skilled labor and a standard technology, by input H , produced by high-skilled workers and modern technologies, and vice versa.

The Y producer minimizes production costs for a given output Y :

$$\begin{aligned} \min C &= w_H H + p_W X_W + p_E X_E \\ \text{s. t. } Y &= H^\alpha (X_W + X_E)^{1-\alpha} \end{aligned} \quad (2)$$

w_H is the price of high-skilled labor intensive input H , p_W and p_E are the prices of X , manufactured in the Western or Eastern region. The final good Y producer is a price taker, he sets the input factor quantities according to the factor prices w_H , p_W and p_E .

The intermediate good is manufactured by independent rivalling firms located in Western Europe or Eastern Europe, respectively. We call the former case outsourcing, the latter case offshore outsourcing. Intermediate good X production is represented by using Cobb-Douglas functions with the inputs capital K and low-skilled labor L with constant returns to scale:

$$X_i = A_i K_i^{\beta_i} L_i^{1-\beta_i}; \quad 0 < \beta_i < 1, \quad i = [W; E] \quad (3)$$

The production processes in West and East differ in technologies A_i and in real wages w_i . Capital is not mobile across borders. Furthermore, differences in the returns to capital investment r_i and in the elasticities of production β_i and $1 - \beta_i$ are possible.

Cost minimizing X manufacturing leads to the following marginal costs c_i :

$$c_i = \beta_i^{-\beta_i} (1 - \beta_i)^{\beta_i - 1} A_i^{-1} r_i^{\beta_i} w_i^{1-\beta_i} \quad (4)$$

Marginal costs c_i are assumed constant and equal to the cost per unit of output. Marginal costs are derived from the exogenous parameters technology A_i , real wage w_i , real return rate on investment r_i and the Cobb-Douglas function exponents β_i and $1 - \beta_i$. If the returns on investment and the exponents are similar in the East and West, a cost advantage can be achieved via a more efficient technology or a lower wage level.

X producers maximize their profits and have oligopolistic (monopolistic) power reflected in the price for X depending on the quantity of $X = X_W + X_E$:

$$\max \Pi_i = p_i(X) \cdot X_i - c_i \cdot X_i \quad (5)$$

Without any market power of X producers, prices would be equal to marginal costs. In general a productivity gap between East and West exists, and hence intermediate good production takes place in the area with lower marginal costs only.

In the case of Bertrand competition (price competition) intermediate good manufacturing occurs only in the cheaper region, too. Now the cheaper producer can increase the price for X and reduce the production quantity of X , but if he increases the price for X more than to the marginal costs of the rival X producer, he will lose all the demand for his product.

Cournot competition (competition in quantities) is the interesting case referring to the current tendencies in the European automotive industry as well as analytically. Under the assumption of pure Cournot competition the Eastern and Western company offer the intermediate good X at the same price $p_X = p_W = p_E$. The firms optimize their supply of X taking the rivals reaction and the demand function for X given by the Western final good Y producer into account. The conditional factor demand function can be derived from (2):

$$X(Y, w_H, p_x) = \left(\frac{\alpha p_x}{(1-\alpha)w_H} \right)^{-\alpha} Y \quad (6)$$

Total demand for X falls with the price p_X and increases with w_H , the price of the skilled labor intensive good H . Solving (6) for p_X yields the inverse factor demand function for X :

$$p_X = \frac{1-\alpha}{\alpha} w_H X^{-\frac{1}{\alpha}} Y^{\frac{1}{\alpha}} \quad (7)$$

Obviously, the intermediate good suppliers face a non-linear and downward-sloping inverse factor demand function with respect to total X . Since $\alpha < 1$, expanding the supply of total X leads to a more than proportional fall in the price p_X . Hence, a monopolist would choose the output as small as possible, but in the oligopoly the situation is different. For every given positive quantity of one supplier there is an optimal output of the rival, which results in an equilibrium with positive quantities.

In market equilibrium supply equals demand for X at the price p_X , so that we can insert (7) into (5):

$$\max \Pi_i = \frac{1-\alpha}{\alpha} w_H X^{-\frac{1}{\alpha}} Y^{\frac{1}{\alpha}} \cdot X_i - c_i \cdot X_i \quad (8)$$

Accordingly, the oligopolists maximize their profits Π_i by choosing their production quantities X_i and by taking the total amount X that includes their own and their rival's quantity into account. This results in the following first order conditions for profit maximization representing the oligopolists reaction functions with $i = [W; E]$:

$$\frac{d\Pi_i}{dX_i} = \frac{1-\alpha}{\alpha} w_H Y^{\frac{1}{\alpha}} \left(-\frac{1}{\alpha} X^{-\frac{1}{\alpha}-1} \cdot X_i + X^{-\frac{1}{\alpha}} \right) - c_i = 0 \quad (9)$$

3 Second order condition for a profit maximum

Now we want to show that the second order condition for the profit maximization problem in equation (8) is indeed fulfilled. We first rewrite the first order condition in equation (9) above using equation (7):

$$\begin{aligned} \frac{\partial \Pi_i}{\partial X_i} &= p_X \left(1 - \frac{1}{\alpha} \frac{X_i}{X} \right) - c_i = 0 \\ \Leftrightarrow 1 - \frac{1}{\alpha} \frac{X_i}{X} &= \frac{c_i}{p_X} \end{aligned} \quad (10)$$

The second order condition for a profit maximum is the derivative of (9) and can be rewritten using (10):

$$\frac{\partial^2 \Pi_i}{\partial X_i^2} = \frac{\partial p_X}{\partial X_i} \left(1 - \frac{1}{\alpha} \frac{X_i}{X} \right) - \frac{1}{\alpha} \frac{p_X}{X} \left(1 - \frac{X_i}{X} \right) < 0 \quad (11)$$

This condition is obviously fulfilled: $\frac{\partial p_X}{\partial X_i}$ is negative according to equation (7) with

$X = \sum^i X_i, i = [W; E]$ and in accordance with basic economic intuition. $1 - \frac{1}{\alpha} \frac{X_i}{X}$ is positive according to (10). Herein, equation (7) ensures that p_X is indeed positive, while marginal costs c_i and α are positive per definition. Finally, in the last term in (11), $1 - \frac{X_i}{X}$ is positive, because all output values are positive and X_i cannot exceed X . (X_i never becomes zero as long as marginal production costs c_i are positive.)

4 Condition for adverse Cournot competition

This section proves analytically that a cost increase in a region reduces output in this region as expected. It then shows mathematically that under certain conditions a cost increase in one region not only decreases output in this region, but additionally slightly decreases output in the other region. Such an adverse Cournot behaviour is a main result of the considerations and simulations in Hübler (2008). This surprising result stems from the interconnection of production in the two regions via Cournot competition.

We recall and rewrite equation (11) in Hübler (2008), where c_E denotes marginal production costs in the East, c_W denotes marginal production costs in the West, and $c_{WE} = c_W / c_E$.

$$V_{WE} \equiv \frac{X_W}{X_E} = \frac{(1-\alpha)c_{WE} + \alpha}{(1-\alpha) + \alpha c_{WE}} = \frac{(1-\alpha)c_W + \alpha c_E}{(1-\alpha)c_E + \alpha c_W} \quad (12)$$

and insert this expression into Hübler (2008), equation (15), second expression:

$$\begin{aligned} X_W &= \left(\alpha' \frac{w_H}{c_W + c_E} \right)^\alpha \frac{1}{1 + \frac{(1-\alpha)c_E + \alpha c_W}{(1-\alpha)c_W + \alpha c_E}} Y \\ &= \Theta \frac{1}{(c_W + c_E)^\alpha} \frac{(1-\alpha)c_W + \alpha c_E}{(1-\alpha)c_W + \alpha c_E + (1-\alpha)c_E + \alpha c_W} \\ &= \Theta \frac{1}{(c_W + c_E)^\alpha} \frac{(1-\alpha)c_W + \alpha c_E}{c_W + c_E} = \Theta \frac{(1-\alpha)c_W + \alpha c_E}{(c_W + c_E)^{1+\alpha}}, \quad \Theta = (\alpha' w_H)^\alpha Y \end{aligned} \quad (13)$$

We first calculate a differential that tells us how production in the West (in absolute terms) reacts to marginal production cost changes in the West (in absolute terms), keeping all other variables constant. (Note that calculating the differential dX_E / dc_E leads to analogue results, i. e. the model is symmetric.)

$$\begin{aligned}
\frac{dX_W}{dc_W} &= \ominus \frac{(1-\alpha)(c_W + c_E)^{1+\alpha} - [(1-\alpha)c_W + \alpha c_E](1+\alpha)(c_W + c_E)^\alpha}{(c_W + c_E)^{2(1+\alpha)}} \\
&= \ominus \frac{(1-\alpha)(c_W + c_E) - [(1-\alpha)c_W + \alpha c_E](1+\alpha)}{(c_W + c_E)^{2+\alpha}} \\
&= \ominus \frac{-\alpha(1-\alpha)c_W - (\alpha^2 + 2\alpha - 1)c_E}{(c_W + c_E)^{2+\alpha}}
\end{aligned} \tag{14}$$

Is this expression smaller or larger than zero? At first, all variables are positive per economically plausible assumption. At second, we assume $\alpha =]0.5; 1[$ throughout the paper. It follows that the numerator is clearly negative, while the denominator is clearly positive. Therefore, the whole differential is negative. As a consequence, higher production costs in the West, for example due to a rising Western wage level, reduce Western production and hence Western employment in absolute terms as expected.

We now derive a differential from (13) that tells us how production in the West (in absolute terms) reacts to marginal production cost changes in the East (in absolute terms), keeping all other variables constant:

$$\begin{aligned}
\frac{dX_W}{dc_E} &= \ominus \frac{\alpha(c_W + c_E)^{1+\alpha} - [(1-\alpha)c_W + \alpha c_E](1+\alpha)(c_W + c_E)^\alpha}{(c_W + c_E)^{2(1+\alpha)}} \\
&= \ominus \frac{\alpha(c_W + c_E) - [(1-\alpha)c_W + \alpha c_E](1+\alpha)}{(c_W + c_E)^{2+\alpha}} \\
&= \ominus \frac{-\alpha^2 c_E + (\alpha^2 + \alpha - 1)c_W}{(c_W + c_E)^{2+\alpha}}
\end{aligned} \tag{15}$$

The condition for the differential to be smaller than zero is:

$$\begin{aligned}
& -\alpha^2 c_E + (\alpha^2 + \alpha - 1)c_W < 0 \\
\Leftrightarrow & \frac{\alpha^2 + \alpha - 1}{\alpha^2} < \frac{c_E}{c_W} =: c_{EW} \tag{16}
\end{aligned}$$

If the condition is fulfilled, a cost increase in the East results in an output decrease in the West, this means there is “adverse Cournot competition”. We first check in which cases this condition is fulfilled for $\alpha =]0.5; 1[$ as assumed throughout Hübler (2008). If α is slightly below one, the whole expression on the left hand side of (16) is also slightly below one. Then condition (16) is still fulfilled when c_E is at least as large as c_W . This leads to the first result: Rising costs in the high-cost region always reduce output in the low-cost region. (Note that calculating the differential dX_E / dc_W leads to analogue results, i. e. the model is symmetric.)

If α is slightly above 0.5 the left hand side of condition (16) is slightly higher than -1. Solving $\alpha^2 + \alpha - 1 = 0$ yields $\alpha_{1,2} = \frac{-1 \pm \sqrt{5}}{2}$. The positive solution is $\alpha_1 \approx 0.618$. This leads to the second result: In the interval $\alpha =]0.5; 0.618[$ rising costs in the low-cost region always reduce output in the high-cost region, independent of the production costs in the East and the West. On the contrary, in the interval $\alpha = [0.618; 1[$ rising costs in the low-cost region can decrease or increase output in the high-cost region depending on the cost differential c_{EW} . This leads to the third result: In the interval $\alpha = [0.618; 1[$ rising costs in the low-cost region reduce output in the high-cost region only if the cost difference between the high- and low-cost region is small (so that c_{EW} is high). If the cost difference is large (so that c_{EW} is small), rising costs in the low-cost region increase output in the high-cost region.

In the simulations in Hübler (2008), section 8, we assumed $\alpha = 0.66$. This leads to a critical cost ratio $c_{EW} \approx 0.219$. If the cost ratio c_{EW} is higher than this critical value within the interval $\alpha = [0.618; 1[$, the differential in (15) becomes negative, otherwise it becomes positive. The cost ratio with respect to average wage levels between the Czech Republic and Germany, for example, is just in the area of this critical value. Hence, according the model, further rising costs in the Czech Republic would likely have a slightly negative impact on German intermediate output. On the other hand, cost increases in a country like Romania would have a positive effect on

German intermediate output, because the cost ratio is larger than the critical value.²

Moreover, α represents the income share of the high-skilled labor and high-tech intensive input factor. This kind of input is used in final production, which is only located in the West. Thus, following the results above, a higher income share of the high-skilled and high-tech intensive factor reduces the likelihood that cost increases in the East are harmful for intermediate production in the West.

However, we need to take into account that in our partial equilibrium analysis the output of the final product is kept constant. Therefore, the analysis does not take into account how cost changes in intermediate production affect the output quantity in final production.

5 Errata

The last sentence in Hübler (2008), page 177 should read: For c_{WE} towards infinity or towards zero the elasticity goes towards zero (Figure 1).

The first equation in (15) in Hübler (2008), page 178 correctly reads:

$$X_E = \left(\alpha' \frac{w_H}{c_W + c_E} \right)^\alpha \frac{1}{1 + V_{WE}} Y$$

6 Conclusion

The analytic treatment of the model set up in Hübler (2008) has provided some new insights. Naturally, rising costs in one region reduce intermediate good output in the same region. The interesting question is under which conditions rising costs in one region also reduce output in the other region to some (possibly small) extent. This is the case of adverse Cournot competition.

The analysis shows that rising costs in the high-cost region always reduce output in the low-cost region. Furthermore, if the factor demand function is relatively steep, rising costs in the low-cost region always reduce output in the high-cost region, independent of the production costs in the East and the West. On the contrary, if the

² Note that according to the model a large cost differential between Romania and Germany implies that the main part of intermediate production is located in Romania.

factor demand function is relatively flat, rising costs in the low-cost region can decrease or increase output in the high-cost region depending on the cost differential between East and West. In this case, rising costs in the low-cost region reduce output in the high-cost region only if the cost difference between the high- and low-cost region is small.

To illustrate the results, suppose the income share of the high-tech and high-skilled labor intensive input factor is relatively high so that the factor demand function is relatively flat. Then according to the model, rising costs in the Czech Republic would likely have a negative impact on German intermediate good output. On the other hand, cost increases in a country like Romania would have a positive effect on German intermediate good output, because the cost ratio is larger than the critical value. Believing this result, policy makers need not be too concerned about reallocation of production to Eastern European countries with very low production costs due to changes in production costs (for example via wage bargaining) within Eastern countries. They should be more concerned about the reallocation effects due to production cost changes in countries where production costs are only slightly lower than in the home country. This is in accordance with the previous outcome of Hübler (2008), stating that changes in relative costs cause larger re-allocations of intermediate production, if the interregional cost difference is smaller.

Finally, following the results, a higher income share of the high-skilled and high-tech intensive factor reduces the likelihood that cost increases in the East are harmful for intermediate production in the West. This implies that a policy that fosters education and innovation as “assets” for production *at home*, makes companies less “vulnerable” to cost changes in other countries when taking offshore outsourcing into account.

However, the analysis says nothing about the magnitude and the economic significance of the results, and the caveats pointed out by Hübler (2008) apply.

References

Hübler, Michael (2008). The Labor Market Effects of Outsourcing Parts and Components: A Simple Model with Cournot Competition. *Aussenwirtschaft* 63(2), 167-194.

Chapter 2:
Avoiding the Trap:
The Dynamic Interaction of North-South Capital Mobility and Technology Diffusion

Forthcoming in The Journal of International Trade & Economic Development,
Taylor & Francis

<http://www.informaworld.com/smpp/title~content=t713722379>

Chapter 2:
Avoiding the Trap: The Dynamic Interaction of
North-South Capital Mobility and Technology Diffusion *

Michael Hübler †

July 23, 2009

Abstract

This paper analyzes a stylized model of international capital mobility and diffusion of embodied technologies from North to South. The South can fall behind in terms of technologies or get trapped in a situation in which it is unable to attract foreign capital and embodied technologies if it is too far away from the technology frontier and if its absorptive capacity is too low. The paper reconciles the view that technological catching up is stronger the larger the technology gap with the alternative view that technological catching up is strongest at a medium technology gap. The closer the South is to the technology frontier the more beneficial is a higher income share of foreign capital.

JEL Classifications: F21, O11, O33, O47

Keywords: Technology diffusion, technology transfer, capital mobility,
FDI, human capital, absorptive capacity

*I thank Johannes Bröcker, an anonymous referee, Fabrizio Zilibotti, Wolfgang Buchholz, Sonja Peterson, Gernot Klepper, Daiju Narita, Zhanna Kapsalyamova and Manfred Wiebelt for very helpful comments and Sorin Krammer for his support.

†Kiel Institute for the World Economy, Düsternbrooker Weg 120, 24105 Kiel, Germany, Email: michael.huebler@ifw-kiel.de, Tel: +49-431-8814-401.

1 Introduction

Technology diffusion from industrialized to developing countries is a promising way of economic development (compare World Bank 2008). Moreover, the potential of international technology diffusion for reducing the energy intensities of economic activities in developing countries and thus for reducing greenhouse gas emissions has become increasingly important. The 2007 Bali conference on climate policy emphasized the responsibility of the industrialized countries to help the developing countries to achieve their emission targets. Most likely, adopting advanced technologies from industrialized countries is much faster and more efficient than inventing own technologies in developing countries. Among the potential channels of international technology diffusion - trade, foreign direct investment, migration, patent citations, scientific literature, communication and information technologies - this paper focuses on international capital movements or in other words foreign direct investment (FDI). Capital flows jointly with embodied technologies to recipient countries, improves their capital and technology endowments and additionally creates technology spillovers to local firms (via imitation of products and machinery, demonstration effects, labor turnover and vertical linkages, increased competition of foreign and domestic firms; compare Saggi 2002).

While some developing countries have been able to converge towards the industrialized countries in terms of technology levels, other developing countries have fallen further behind (World Bank 2008). Herein, the human capital endowment of a recipient developing country is a main determinant of the success of technology adoption (see for instance World Bank 1993). Additionally, own innovation performed in a developing country might be an effective substitute or support for technology adoption (Lall and Urata 2003).

While there is a broad strand of empirical literature on the effects of capital inflows on productivity and growth, theories on technology diffusion via capital inflows are rare (for example Findlay 1978, Wang and Blomström 1992, Das 1987, Mayer-Foulkes and Nunnenkamp 2009). And there is no common intuitive theory that describes the dynamic interactions of international capital movements, technology diffusion, innovation and absorptive capacities of recipient countries.

Our study therefore investigates the dynamic interaction of North-South capital mobility and technology diffusion building on Nelson and Phelps (1966, in the following denoted by N&P). Our study deals with the question, whether market forces enable international transfer of capital and embodied technologies in such a way that the technology

growth rate of a developing country catches up with the growth rate of an industrialized country so that the technology gap narrows. If there is an automatic technology diffusion mechanism that narrows the technology gap, this will be an argument for development and climate policy not to overreact and to let market forces work instead. If, on the other hand, there is a risk that technology diffusion does not occur sufficiently based on market forces, there will be need for policy interventions that actively support technology diffusion processes and that improve the absorptive capacities of developing countries.

The main contribution of this paper is to provide a theoretical explanation for a phenomenon frequently identified in the empirical literature and for example discussed by the World Bank (2008): The North-South technology gap widens and convergence of technology growth rates fails in some developing countries despite increasing global economic integration via international investment and other channels. Our analysis therefore provides one possible explanation (among others for example described by Hanson 2001) for the diverse results of the econometric literature about the impact of FDI on growth and the role of human capital for technology diffusion. A consecutive question is, whether own innovation in developing countries can remedy such a convergence failure.

Furthermore, there is the point of view that technology diffusion is strongest at a medium technology gap as described by Benhabib and Spiegel (2005). So far, this has been an artificial construction in order to explain the observation that certain countries fall behind in terms of technologies. A contribution of this paper is to derive this outcome theoretically through the introduction of international capital mobility (for the special case of technology diffusion via capital mobility).

We also examine the potential of own innovation in the South for narrowing the technology gap and for preventing a convergence failure. Although technology diffusion and own innovation in the South are basically substitutes, they can positively interact as complements in the short-run. The reason is that innovations raise total factor productivity, which in turn attracts more foreign capital and embodied technologies. On the other hand, own innovation in the South cannot prevent falling behind in terms of technologies in the long-run, except when the South is as innovative as the North. Different to the original theory by N&P, in our model a higher initial technology *level* (innovative capability) in the North increases the technology gap in the presence of own innovation in the South, because performing own innovation reduces the North-South technology gap and hence the possibility to adopt foreign technologies via international technology diffusion.

According to our model, a larger income share of internationally mobile foreign capital is more beneficial for technology diffusion when the South is close to the technology frontier.

The paper proceeds as follows. Section 2 briefly reviews the related theoretical and empirical literature. Section 3 re-interprets the N&P theory in the context of FDI in general form. Section 4 introduces international capital mobility explicitly in a myopic market solution approach. Section 5 critically discusses the results of the paper. Section 6 concludes.

2 Literature Background

This section at first describes empirical studies on FDI induced productivity gains and then presents theoretical approaches dealing with productivity gains via international capital mobility.

There is a broad strand of empirical research on international technology diffusion via FDI including cross-country panel analyses as well as case studies for specific countries. Numerous econometric studies examine the effects of FDI inflows on GDP growth of recipient countries or productivity spillovers from foreign to domestic firms - with diverse results. (Kokko 1992, Blomström and Kokko 1998, OECD 2002, Saggi 2002, Keller 2004 and Hoekman and Javorcik 2006 provide detailed literature surveys. Chen and Dunning 1994 also include comprehensive theoretical background information. Concerning East Asian economies see Lall and Urata 2003.) A number of studies confirm the positive effect of human capital on technology diffusion (Benhabib and Spiegel 1994, Crispolti and Marconi 2005, Kneller 2005, Girma 2005, Lai et al. 2006), while others do not confirm it (Sjöholm 1997, Xu and Wang 2000). Some papers additionally find a minimum human capital level which is necessary to enable technological catching up (Borensztein et al. 1998, Crespo et al. 2004, Benhabib and Spiegel 2005, Ciruelos and Wang 2005, also see OECD 2002). Mayer-Foulkes and Nunnenkamp (2009) observe that FDI accelerates economic convergence among high-income countries, while it widens the income differential between the USA and low- and middle-income countries.

Moreover, the econometric literature examines the role of the technology gap between the technology in practice in the recipient country and the technology frontier. Some studies find evidence for the hypothesis that the technology diffusion strength increases the larger the technology gap (Griffith et al. 2002, Girma 2005, Griffith et al. 2004,

weak evidence by Kokko et al. 1996). Others support the hypothesis of an inverted U-shaped relation of the technology gap and the technology diffusion strength (Benhabib and Spiegel 2005, Girma et al. 2001) or a U-shaped relation (Girma and Görg 2007).

The World Bank (2008) summarizes the effects of technology diffusion to developing countries as follows:

”The level of technological achievement in developing countries has converged with that of high-income countries over the past 15 years. A sustained policy of increased openness to foreign trade and foreign direct investment (FDI), plus increased investments in human capital, have contributed to substantial improvements in technological achievement in developing countries over the past 15 years. And despite rapid progress at the technological frontier, technological achievement in both low- and middle-income countries has increased much more rapidly than in high-income countries. As a result, developing countries have closed the relative gap with high-income countries. However, the gap remains large. Moreover, the strong aggregate performance of low-income countries reflects large improvements in technological achievement by some, but much more modest advances by the majority. As a consequence, many are only maintaining pace with, or even losing ground to, high-income countries.”

An important strand of the theoretical literature deals with endogenous growth via horizontal or vertical product (variety) improvements, for example Krugman (1979), Romer (1990), Grossman and Helpman (1991), Barro and Sala-i-Martin (1997), Aghion and Howitt (2005). Acemoglu, Aghion and Zilibotti (2003a, 2003b) provide full micro founded analyses of imitation and innovation dependent on the distance to the technology frontier.

Our study abstracts from the sources of innovation and economic activities on the micro level and rather examines technology diffusion processes on a macro level in order to understand the large-scale time paths and interactions. We build on Nelson and Phelps (1966, henceforth denoted by N&P). In their macroeconomic model, N&P formalize the so-called Veblen-Gerschenkron effect.¹ Their intuitive approach has been applied in a number of studies.²

Only few theoretical models specifically deal with FDI as a channel of technology transfer. Findlay (1978) sets up a model of technology diffusion through FDI, where the rate of technical progress in the recipient backward region is a negative function of the technology level in the backward region relative to the level in the advanced region, and a positive function of the stock ratio of foreign to domestic capital within the

¹Gerschenkron (1962) studies the phenomenon of economic catching up of countries that have fallen behind.

²Aghion (2007), for instance, applies the N&P approach to examine the effect of education on growth.

backward region. Das (1987) examines the optimal dynamic behavior of multinational firms when knowledge spillovers to rivals in the host country occur. He shows that the indigenous firms do not necessarily benefit from the technology transfer to subsidiaries of multinationals firms, while the host economy as a whole always benefits. Wang and Blomström (1992) endogenize technology transfer via capital mobility including costs of transferring technologies and of learning. Building on N&P, Diao et al. (2005) multiply the share of intermediate goods imports in GDP by the share of imported capital in GDP and by the distance to the technology frontier. Their approach takes into account that foreign firms likely import advanced capital goods (from their home countries). Mayer-Foulkes and Nunnenkamp (2009) show which conditions concerning the strength of technology transfer and convergence forces must be fulfilled in order to guarantee international convergence of economic growth rates.

Nonetheless, the existing theories do not directly explain the empirical facts described before. Against this background, the contribution of this paper is to provide an intuitive theoretical description of international technology transfer via capital mobility that provides one possible reason (among other reasons) for the mixed empirical evidence described above.

3 The Nelson and Phelps Theory in the Context of International Capital Mobility

The considerations of N&P are based on the Veblen-Gerschenkron effect (Gerschenkron 1962). According to the N&P theory, technological catching up is faster the larger the gap between the technology in practice and the technology frontier and the better the educational attainment. When the technology level of the "learning" country is low, most of the newly arriving technologies are not yet known and therefore beneficial. The higher the level, which the "learning" country has reached, the more newly appearing technologies are already known and therefore without an additional benefit. Human capital (educational attainment) enhances the technology diffusion speed for every given technology gap, since it improves the ability to adopt and apply new technologies. In case of exogenous technological progress of the frontier, the technology in practice follows the frontier with the same rate of technological progress and with a constant relative technology gap.

The N&P theory can be applied to an industrialized country (technological leader,

denoted by North) that creates the leading technologies and a developing country (technological follower, denoted by South) that adopts technologies from the North and follows with a certain technology gap (compare Benhabib and Spiegel 2005). While in N&P the technology diffusion speed only depends on educational attainment, in our re-interpretation technology diffusion additionally depends on the volume of foreign capital in the South, similar to Findlay (1978). In this context, we re-interpret educational attainment as the absorptive capacity, including all factors that determine the ability of host countries to absorb transferred technologies and to benefit from them.

In contrast to the original N&P theory, empirical evidence shows that many developing countries are not able to catch up. For that reason, the original equation by N&P has been artificially modified in the literature so that catching up is fastest at a medium technology gap and decreases the larger the gap and also the smaller the gap (logistic model, e.g. described by Benhabib and Spiegel 2005). As a consequence, it is possible that convergence fails so that a country falls further behind in terms of technology levels. But this assumption seems intuitively not convincing. Why should a country with high educational attainment, a stable political and legal system and good infrastructure not be able to catch up in a certain sector or field of technology such as solar energy generation where it has completely missed the newest technological development? The inability of developing countries to catch up probably rather lies in the determinants of technology accumulation like education and the legal system, not in the low technology level itself (compare OECD 2002). If these determinants of technology diffusion are sufficiently present, technological catching up is possible even far away from the frontier. Our considerations therefore follow this point of view.

Van Meijl and van Tongeren (1999) assume that international spillovers from trade are quantitatively higher when countries are similar in their economic structure. This view is in accordance with our model. Given a high education level, infrastructure etc. in the leading industrialized country, the improvement of these factors in the developing country makes it more similar to the developed country and thus increases the technology spillover strength.

The first section 3.1 explains how technological catching up of the South is influenced by the rate of technological progress in the North and by the absorptive capacity of the South based on N&P. Section 3.2 introduces own innovation in the South and investigates its interaction with technology diffusion. 3.3 briefly explains the allocation of internationally mobile capital in the long-run.

3.1 A Re-Interpretation of the Nelson and Phelps Theory

The equations below describe the basic model formulated by N&P. Our extension is the introduction of capital as a determinant of the technology diffusion speed. Throughout the paper, n denotes the North, and s denotes the South.³

$$\dot{A}_s = \phi_s(A_n - A_s), \quad \phi_s = \phi_s(H, K) \quad (1)$$

$$A_n = A_n(0)e^{\lambda t} \quad (2)$$

A_s is the endogenous technology level in practice in the destination country, the South. A_n is the level of the exogenous technology frontier in the North. We assume that capital transferred from the North to the South embodies technologies up to this frontier level. A_n and A_s can be interpreted as knowledge capital stocks that are accumulated and that determine total factor productivity. Technology diffusion increases total factor productivity in the South A_s . λ is the exogenous growth rate of the technology frontier. $A_n(0)$ is the level of the frontier at $t = 0$. ϕ_s is the (in this section completely exogenous) spillover strength including the imitation capability of the South and has the following properties:

$$\frac{\partial \phi_s(H, K)}{\partial H} > 0, \quad \frac{\partial \phi_s(H, K)}{\partial K} > 0$$

ϕ_s is an increasing function of the relative human capital (educational attainment) $H = \frac{H_s}{H_n}$, i.e. the human capital level of the South relative to the human capital level of the North. Throughout the paper, we interpret H in a broader sense as the absorptive capacity including infrastructure, the legal and political framework etc. of the recipient country. A specification in relative terms seems reasonable, because technologies are invented in the North given the North's educational level and absorbed by the South given the South's educational level. The chance to absorb Northern technologies successfully, increases when the South's educational level comes closer to the North's level. H also incorporates naturally given factors that influence technology diffusion. H increases ceteris paribus with the spatial concentration of economic activities, because spatial concentration eases productivity spillovers in accordance with agglomeration theory. China, for example, has established Special Economic Zones in order to concentrate foreign economic activities and to maximize spillovers. Also, small economic areas with

³ A_s and A_n are time dependent variables. Time indices of variables are not shown explicitly. Time derivatives are denoted by dots.

high population densities like Hong Kong and Taiwan have shown amazing catching up and economic growth performances. H furthermore captures natural conditions. It decreases if the land is mountainous and landlocked and hence difficult to access, if the climatic conditions are problematic and so forth. We may think of sub-Saharan African countries facing severe detrimental political and natural conditions that prevent the transfer of capital and technologies. H may also increase with the size of the labor force, since the pool of workers suitable for employment in multinational enterprises increases with its size. All kinds of policies that improve technology diffusion can be modeled by raising H . For example, the establishment of China's Special Economic Zones, the Chinese policy to enforce joint ventures of foreign investors with domestic firms, and the Chinese policy to grant foreign investors privileges of various forms.

Compared to N&P, we additionally assume that ϕ_s increases with the South-North ratio of internationally mobile (high-tech) capital $K = \frac{K_s}{K_n}$. This approach is a modification of Findlay's (1978) model. The underlying assumption is that high-tech capital is built up in the North and embodies advanced technologies that have been invented in the North. Capital and embodied technologies are simultaneously transferred to the South. But the technologies are not immediately available in all Southern production processes. They rather need time to diffuse into and through the Southern economy (via product and process imitation, learning from foreign managers, engineers or workers, vertical and horizontal linkages between suppliers and customers, productivity gains through increased competition etc.). The quality of the technologies embodied in foreign capital is supposed to rise at the constant rate A_n over time due to Northern innovation activities. We meanwhile assume K to be exogenously given at each certain point of time.

H and K act as complements, i.e. they enhance each other.⁴ More foreign investment potentially yields even larger positive spillovers when at the same time the absorptive capacity is higher, in accordance with the empirical findings. Technology diffusion ceases if H or K are zero.

Without technological progress of the technological frontier, the technology level of the developing country catches up completely with the frontier. Following N&P, we rather assume exogenous exponential technological progress of the frontier and reinterpret their results with respect to K . Like N&P, we first solve differential equation (1) and then calculate the relative long-run technology gap. The conclusions of N&P con-

⁴This is satisfied in a multiplicative specification that we will use in section 4, but not in an additive specification.

cerning the effect of H on technology diffusion can then be directly transferred to the effect of K .⁵

$$\frac{A_n - A_s}{A_s} = \frac{\lambda}{\phi_s} \quad (3)$$

$$A := \frac{A_s}{A_n} = \frac{\phi_s}{\phi_s + \lambda} = \frac{1}{1 + \frac{\lambda}{\phi_s}} < 1 \quad (4)$$

Equation (4) shows that in the long-run the technology level of the South has a constant ratio A to the technology level of the North, when K and H stay constant. Since ϕ_s is an increasing function of K , a higher foreign capital intensity K and a higher human capital endowment H reduce the relative equilibrium technology gap as shown in equation (4).

The elasticity of the Southern technology level in the steady state with respect to the foreign capital intensity reads:⁶

$$\frac{\partial A_s}{\partial K} \frac{K}{A_s} = \frac{\frac{\partial \phi_s}{\partial K}}{\phi_s} \frac{K \lambda}{\phi_s + \lambda} \quad (5)$$

According to (5), the relative increase in the technology level due to a relative increase in foreign capital is greater the higher the technological progress λ of the frontier. A subsidy on internationally mobile foreign capital can be economically justified because of the positive technology spillover of mobile capital (based on the stylized narrow view of this model). A subsidy increases the foreign capital share. It follows from (5) that the benefit of the subsidy is greater when technological progress in the North is higher.

3.2 Innovation in the South

According to equations (3) and (4), the technology gap cannot be completely closed through technology diffusion as long as H and K are finite. A possible remedy is to add own innovation in the South as described by Benhabib and Spiegel (2005). Our contribution is to solve the differential equation including innovation in the South.⁷ We

⁵For more detailed calculations see N&P and our calculations in equations 6 to 10, which are a generalized form of the following basic calculations.

⁶ A_s is given by equation (9) with θ_s set to zero.

⁷Our basic relation implies imperfect technology diffusion. Perfect technology diffusion, i.e. $\phi_s \rightarrow \infty$, would make own innovation in the South superfluous. For a further discussion see Grossman and Helpman (1991).

set up a diffusion equation similar to Benhabib and Spiegel (2005):

$$\dot{A}_s = \phi_s(A_n - A_s) + \theta_s \gamma e^{\gamma t} \quad (6)$$

$$\phi_s = \phi_s(H, K), \quad \theta_s = \theta_s(H_s, R_s), \quad \gamma = \gamma(H_s, R_s) \quad (7)$$

We assume an exogenous Southern innovation rate γ which increases with the input of R&D resources R_s and with H_s .⁸ θ_s is the Southern innovation level in $t = 0$, which can be interpreted as the accumulated knowledge up to this point of time. In this sense, it depends on the values of H_s and R_s before $t = 0$. Similar to ϕ_s , θ_s determines how successfully new technologies can be implemented into production processes in the South. Human capital, infrastructure etc. improve ϕ_s , the technology diffusion capability, as well as θ_s which we call in the analog way innovative capability. Herein, we make a simplifying assumption: The time lag between the appearance of new technologies and their full implementation in production is relatively small in case of Southern innovations, because they are directly created for local production and for the local market and fit to the local abilities and circumstances. Foreign technologies on the other hand, often exceed the existing local abilities by far and do not directly fit to the local circumstances. To make the argument more illustrative: Sub-Saharan African countries might develop a new water extraction and storage system through public funding that will be used almost everywhere within a couple of years and that will create a welfare gain. But they will certainly not be able to adopt all relevant chemical know-how from US firms which would be necessary to produce the same advanced pharmaceuticals autonomously, although it would certainly boost African welfare.

The solution of differential equation (6) becomes (like in N&P plus the additional source of technological progress):⁹

$$A_s = \left[A_s(0) - \frac{\phi_s A_n(0)}{\phi_s + \lambda} - \frac{\gamma \theta_s}{\phi_s + \gamma} \right] e^{-\phi_s t} + \frac{\phi_s A_n(0)}{\phi_s + \lambda} e^{\lambda t} + \frac{\gamma \theta_s}{\phi_s + \gamma} e^{\gamma t} \quad (8)$$

The first term describes the transition process and vanishes as $t \rightarrow \infty$. In accordance with Benhabib and Spiegel (2005), the leader and the follower still grow at the same *rate* and a certain relative gap in the long-run, when the additional technological progress term is added. Complete catching up of the South's technology level to the North's

⁸Here we write indices s , because the influence of human capital, infrastructure, R&D etc. in the South on Southern innovation is independent of the corresponding Northern values.

⁹Note that the start value $A_s(0)$ includes knowledge gained from technology diffusion as well as from own Southern innovation.

technology *level* is only possible, if the South's own innovative capability is (at least) identical to the North's innovation capability, i.e. $\gamma = \lambda$ and additionally $\theta_s = A_n(0)$. But this case occurs unlikely in reality, since the innovative capability is lower in developing countries than in industrialized countries. However, own innovation in the South has the potential to accelerate the convergence process, in other words, it has the potential to lift up the South to a higher transition path.

Different to the literature, we now assume $\gamma = \lambda$, i.e. identical innovation rates in the North and in the South so that θ_s becomes the policy parameter controlling the innovation strength, since in general $\theta_s \neq A_n(0)$.¹⁰ We can then derive the long-run technology level of the South:

$$A_s = \frac{\phi_s A_n(0) + \lambda \theta_s}{\phi_s + \lambda} e^{\lambda t} \quad (9)$$

We express the long-run technology ratio A , now including the South's own innovation level, as:

$$A = \frac{\phi_s A_n(0) + \theta_s \lambda}{\phi_s A_n(0) + A_n(0) \lambda} = \frac{\frac{1}{\lambda} + \frac{\theta_s}{\phi_s A_n(0)}}{\frac{1}{\lambda} + \frac{1}{\phi_s}} < 1 \quad (10)$$

This equation clearly shows that complete catching up is possible in two ways: At first, the spillover strength ϕ_s becomes infinitely strong (due to an infinitely large absorptive capacity or an infinitely high foreign capital intensity). At second, the Southern innovative capability θ_s becomes equal to the leading innovative capability $A_n(0)$. Both ways seem not realistic, the latter at least for developing countries. Furthermore, when the rate of technological progress (which is assumed to be equal in the North and South) increases, the relative technology gap will mainly be determined by the South's innovative capability θ_s relative to the North's innovative capability given by $A_n(0)$. According to the original N&P model, the level $A_n(0)$ of the technology frontier does not influence the relative technology gap, while a higher rate of technological progress of the frontier increases the gap. Our model additionally considers own innovation in the South. Now, in the presence of innovation, a higher initial technology *level* of the frontier $A_n(0)$, i.e. a better innovative capability of the North, widens the gap. The intuition is the following: In our model framework a higher level of Northern technologies enlarges the North-South technology gap which in turn enhances the technology diffusion speed. On the

¹⁰While in the former analysis a higher technological progress λ of the technological leader made technology adoption more effective, the influence of λ on the technologies used in the South is now ambiguous, because innovation in the South is assumed to grow with λ , too.

other hand, own Southern innovations reduce the North-South technology gap, which in turn reduces international technology diffusion. For this reason, the South cannot fully benefit from a higher Northern technology level in the presence of own Southern innovation. The consequence is, that the North fully benefits from its new innovations, while the South does not. Thus, the relative gap increases with $A_n(0)$ in the presence of own innovation in the South.

While both alternatives of complete catching up seem not realistic, we now examine the effectiveness of investing into a higher innovative capability θ_s . Technology transfer and innovation are always beneficial from the point of view of the model, but the effectiveness of certain policy measures depends on the other model parameters. Following N&P, we derive a technology elasticity, in our case with respect to innovation in the South. This allows us to investigate the effectiveness of enhancing Southern innovation for narrowing the North-South technology gap:

$$\frac{\partial A_s}{\partial \theta_s} \frac{\theta_s}{A_s} = \frac{1}{\frac{\phi_s A_s(0)}{\lambda \theta_s} + 1} \quad (11)$$

According to (11), the relative change in the Southern technology level due to a relative change in the Southern innovative capability (around a certain value of the innovative capability) is higher when ϕ_s and $A_s(0)$ are low. We recall that ϕ_s is an increasing function of the absorptive capacity and of the foreign capital intensity. The intuition is similar to what we considered before: Own Southern innovations increase the Southern technology level and decrease at the same time the North-South technology gap so that international technology diffusion slows down. In this sense, technology diffusion and own innovation act as substitutes.

Now we relax the assumption that Northern and Southern innovation processes have the same *rate* of progress again. In general, it is plausible to assume that the Southern innovation process has a lower rate than the Northern frontier process, i.e. $\gamma < \lambda$. We divide equation (6) by A_s to obtain the growth rate of the Southern technology level $\hat{A}_s = \frac{\dot{A}_s}{A_s}$, where $A = \frac{A_s}{A_n}$:

$$\hat{A}_s = \phi_s \left(\frac{1}{A} - 1 \right) + \frac{\theta_s \gamma e^{\gamma t}}{A_s} \quad (12)$$

A_s grows with the same rate λ as the technology frontier in case of convergence of growth rates, while own innovations grow at a lower rate γ per plausible assumption. Hence, the last term in the equation above will vanish over time, so that the benefit of Southern

innovation vanishes. Nevertheless, innovation in the South accelerates the convergence process to the steady state.

However, the optimal decision on investment in enhancing technology diffusion needs to include the related costs which are neglected in our analysis.

3.3 International Factor Allocation in the Steady State

We now turn to the question how technology diffusion affects the allocation of internationally mobile capital between North and South in the long-run steady state when all adjustment processes are completed. Does the international capital allocation still change in the steady state after the catching up process has been completed?¹¹ The answer is no, which can easily be seen. In the ideal case (neglecting transaction costs) the marginal products of mobile capital and hence their prices are equal in the North and in the South (factor price equalization). If there are transaction costs, there will be a constant difference between the marginal products. The N&P equation yields equal growth rates λ of the technology of North and the South and a constant relative technology ratio A in the steady state.

Hence, with or without transaction costs, in the long-run total factor productivities and marginal products of capital in the North and in the South have a constant ratio, which makes any adjustments of mobile capital stocks economically superfluous (as long as there is no external shock). Moreover, the higher the technology level of the South relative to the technology level of the North, the higher the quantity of mobile capital allocated to the South relative to the quantity of capital allocated to the North.

4 Endogenous International Capital Mobility

While the foreign capital intensity was exogenous in the last section, it is now endogenized in a simple straightforward way: Marginal products of internationally mobile high-tech capital are equalized across North and South.¹² The question is how capital mobility and technology diffusion through capital mobility interact. There is potentially a positive feedback mechanism: A better absorptive capacity and a higher foreign cap-

¹¹Own innovation in the South may or may not be present at any rate $\gamma \leq \lambda$.

¹²The following outcomes also hold when including transaction costs of capital movements. In this case there is a constant difference between the marginal products in the steady state.

ital intensity enhance the technology diffusion speed. This in turn raises the marginal product of internationally mobile capital and thus attracts more mobile capital and so forth. On the other hand, a situation of a low foreign capital endowment and a low absorptive capacity in the South results in a slow technology diffusion speed. This in turn increases the technology gap so that even less capital is allocated to the South. Hence, the South might get trapped concerning its technological development and foreign capital accumulation if the technology diffusion speed does not increase sufficiently far away from the technology frontier.

This section examines these questions. Subsection 4.1 describes the model setup, subsection 4.2 interprets the model.

4.1 The Model

In the following stylized model, the international allocation of capital is purely driven by differences in marginal products of capital without perfect foresight and without internalizing the social benefit of technology transfer. Therefore, our model can be called a myopic market solution approach. So far, a logistic function approach has been used in the literature to model a slower diffusion speed when the technology gap is either small or large (Benhabib and Spiegel 2005). This assumption implies an inverted U-shaped relation between the technology gap and the technology diffusion speed without a direct theoretical explanation. In our model international capital allocation is the explicit mechanism that leads to such a relationship.

We first write the technology diffusion equation (6) without innovation in the South, explicitly including the influence of the absorptive capacity and of foreign capital on technology diffusion:

$$\dot{A}_s = \phi_s(A_n - A_s) \quad (13)$$

$$\phi_s = H^\mu K^\nu \quad (14)$$

H is the absorptive capacity that rises with the Southern human capital endowment, infrastructure, labor endowment and other economically beneficial factors and decreases with economically detrimental factors such as missing access to the sea. H is measured as a South-North ratio, so that a higher similarity between the South and the North enhances the chance to transfer technologies successfully. (Compare section 3.1.) As before, international technology diffusion is driven by *foreign* capital (FDI, multina-

tional enterprises) building on the broad empirical literature. Also as before, a higher South-North ratio of foreign capital K induces a faster technology diffusion when there is a better absorptive capacity H . Therefore, K and H act as complements. This multiplicative specification follows the empirical literature applying interaction terms (e.g. Borensztein et al. 1998, Benhabib and Spiegel 2005, Girma and Görg 2007). It seems reasonable to assume $\mu + \nu \leq 1$ i.e. a decreasing or constant returns to scale specification.

In the next step we derive K from profit maximization of firms under perfect competition and perfect capital mobility, given a Cobb-Douglas type production function. This constant returns to scale production function encompasses total factor productivity A_i , foreign (high-tech) capital K_i , domestic (low-tech) capital D_i and the aggregate input H_i . The additive combination of foreign and domestic capital in the production function implies the co-existence of foreign firms (multinational enterprises) and domestic firms that produce the same output Y_i . H_i encompasses human capital and infrastructure, labor, land and other factors as described before. All factors that enhance (hinder) technology diffusion probably also enhance (hinder) production.

$$Y_i = A_i(K_i^\alpha + D_i^\alpha)H_i^{1-\alpha} = (K_i^\alpha + D_i^\alpha)(\tilde{A}_i H_i)^{1-\alpha} \quad (15)$$

Herein, we use the notation $i = n, s$ and $A_i = \tilde{A}_i^{1-\alpha}$. High-tech capital K_i is built up in the North and perfectly mobile between North and South. It is assumed to stay constant so that $K_n + K_s = 1$. This assumption implies that high-tech capital is a scarce resource.¹³ The marginal product of K_i can be derived as follows:

$$\frac{\partial Y_i}{\partial K_i} = \alpha K_i^{\alpha-1}(\tilde{A}_i H_i)^{1-\alpha} \quad (16)$$

Perfect international mobility of high-tech capital equates the marginal products of K_i :

$$\begin{aligned} \alpha K_n^{\alpha-1}(\tilde{A}_n H_n)^{1-\alpha} &= \alpha K_s^{\alpha-1}(\tilde{A}_s H_s)^{1-\alpha} \\ \Leftrightarrow k_n &= k_s \end{aligned} \quad (17)$$

$$\Rightarrow K = \tilde{A}H \quad (18)$$

¹³Note that in this model the transfer of capital is beneficial both for the North and the South, because in the initial situation the mobile capital earns a higher return in the South than in the North and the return on foreign direct investment is transferred back to the North. Otherwise, no capital would be transferred from North to South.

In the second equation we write K_i in efficiency units $k_i = \frac{K_i}{\tilde{A}_i H_i}$. In the last equation K , \tilde{A} and H all denote South-North ratios.

Domestic (low-tech) capital is assumed to be internationally immobile and without any influence on international technology diffusion. Different to K_i , D_i is accumulated in the North and in the South via saving and investing part of income Y_i . (More details are presented in the Appendix.) In case of closed economies, K_s is zero. In this case, our model coincides with the standard Solow (or alternatively Ramsey) model of a closed economy, and the South grows due to own (exogenous) innovation and due to accumulation of capital D_s . (Own innovation in the South was discussed in section 3.2.) In the following analysis we basically assume that the South rather follows a purely imitation based strategy, i.e. technological progress in the South is dominated by technology diffusion from the North via the transfer of high-tech capital.

Now we simply insert equation (18) with $\tilde{A} = A^{\frac{1}{1-\alpha}}$ into (13) and (14) to obtain:

$$\dot{A}_s = H^{\mu+\nu} A^{\frac{\nu}{1-\alpha}} (A_n - A_s) \quad (19)$$

Equation (21) can be re-written in terms of the growth rate $\hat{A}_s = \frac{\dot{A}_s}{A_s}$.¹⁴

$$\hat{A}_s = H^{\mu+\nu} \underbrace{A^{\frac{\nu}{1-\alpha}}}_{\text{density}} \underbrace{(A^{-1} - 1)}_{\text{distance}} \quad (20)$$

$$\Leftrightarrow \hat{A}_s = H^{\mu+\nu} \left(A^{\frac{\alpha+\nu-1}{1-\alpha}} - A^{\frac{\nu}{1-\alpha}} \right) \quad (21)$$

4.2 Interpretation

This section interprets equations (20) and (21) derived above. Obviously, a higher H is always beneficial for technology diffusion, since it enhances the diffusion speed as well as output per assumption. Both effects enhance each other in a complementary way. In an analog way, a widespread, unaccessible area with disadvantageous natural and climatic conditions and other economically detrimental factors always impede technology diffusion.

Our main aspect under scrutiny is technological catching up and convergence of technology growth rates. Equation (20) decomposes the impact of the technology gap between A_n and A_s into two opposing parts: The first "A term" means that a lower ratio of A_s to A_n reduces the relative amount of foreign capital in the South. It describes

¹⁴The variable A has a maximal value of one, and the exponent of the first "A term" in parentheses in equation (21) is always smaller than the exponent of the second "A term" in parentheses so that the whole expression never becomes negative.

that more foreign capital is allocated to a region if its relative technology level is higher so that technology diffusion is also higher. Let us call this *density effect*. At the same time, the second "A term" in equation (20) states that a larger technology gap increases the possibilities to adopt new technologies so that the technology diffusion speed rises. Let us call this *distance effect*.

It is a priori not clear, which effect dominates. The outcome depends on the values of α and ν . Suppose, the technology growth rate of the South \hat{A}_s is lower than the technology growth rate of the North \hat{A}_n in the initial situation. If the distance effect dominates, the South will fall behind in terms of the technology level until the distance to the technology frontier is so large (i.e. A is so small) that the South's technology growth rate becomes equal to the North's growth rate. In other words, the technology gap widens, but growth rates finally converge. If the density effect dominates, \hat{A}_s steadily decreases with a lower A . In this case, there is no automatic convergence mechanism. Without own innovation in the South and without policy intervention, the South falls increasingly behind in terms of technologies, and the amount of foreign high-tech capital allocated to the South asymptotically drops towards zero. - The South is trapped.

In order to find the condition for the possibility of a convergence failure, we note that in (21) the second term in parentheses vanishes, if the technology ratio A becomes smaller due to a difference in the technology growth rates, $\hat{A}_s < \hat{A}_n$. If the first term in parentheses also decreases in A , the South will be trapped. This situation can only occur when $\nu + \alpha > 1$. This is more likely the case, when α and ν are large, because then the foreign capital endowment in the South reacts more strongly to changes in the technology level A_s . In this case, lowering A_s shifts away a relatively large amount of foreign capital. In case of a linear diffusion model with $\nu = 1$, the condition for the possibility of a growth trap is always fulfilled.

We can now derive another interesting aspect by plotting the function expressed by equations (20) and (21) for different values of α as shown in Figure 2 in the Appendix: The larger the exponent α , which means the income share of foreign capital,¹⁵ the closer the point of maximal diffusion speed is to the technology frontier and the lower the value of A_s is in the maximum. Therefore, a higher income share of foreign capital becomes more beneficial the closer the economy is to the technology frontier. Farther away from the technology frontier, a lower income share of foreign capital becomes more beneficial.

¹⁵For reasons of mathematical simplicity, we have also chosen the income share of domestic low-tech capital to be α . But what matters for this outcome is the income share of foreign high-tech capital, not the income share of domestic low-tech capital. Hence, the result would also hold for different exponents of K_i and D_i .

This outcome stems from the fact, that the demand for foreign capital in the South as a function of A_i described by equation (16) is steeper the higher α . Consequently, when $A = \frac{A_s}{A_n}$ is large, a higher value of α attracts more foreign capital to the South. When A is small so that the South is far away from the technology frontier, a higher value of α attracts less capital to the South. This result is in accordance with what we found before: The likelihood of a convergence failure is higher when α is larger. The reason is that a higher value of α leads to a stronger reaction of the international capital allocation driven by changes in the South-North technology ratio. Thus, the density effect becomes stronger for a higher value of α , and more mobile capital is shifted away from the South when the South falls behind in terms of technologies. Also note that according to Figure 2, increasing α above 0.5 raises \hat{A}_s only slightly. That means, there seems to be a small additional benefit for technology diffusion of an income share of foreign capital higher than 0.5.

We now set $\nu = 1$ and $\alpha = \frac{1}{2}$, which yields an illustrative quadratic form:

$$\hat{A}_s = H^{\mu+1} (A - A^2) \quad (22)$$

Note that the following considerations qualitatively also hold for other values of α . Figure (1) is the (qualitative) graphical representation of equation (22). On the right hand side of Figure 1 the South's technology level is close to the technology frontier given by the North, i.e. A is close to one. On the left hand side it is far away from the technology frontier. The vertical axis shows the technology growth rate of the South \hat{A}_s as a function of the technology ratio A on the horizontal axis. Obviously, the maximal diffusion speed is reached at half the distance to the technology frontier. (This outcome changes, when another value for the income share of foreign capital α is chosen as discussed before. Compare Appendix, Figure 2.)

We notice that starting at a high level of A on the right hand side of Figure 1 in area B_1 , the Southern rate of technological progress \hat{A}_s increases, while the technology ratio A decreases. (We always move on the parabola.) The technology gap widens. In the steady state the technology growth rate of the South has converged to that of the North. Full convergence of growth rates as well as of technology levels implies $\lim_{t \rightarrow \infty} A = 1$. This case does not occur in this model with exogenous exponential technological progress. It occurs however, if the technology frontier stays constant. (Then λ is zero in equation 4 resulting in $A = 1$.)

We then start at a medium distance to the technology frontier in the middle of Figure

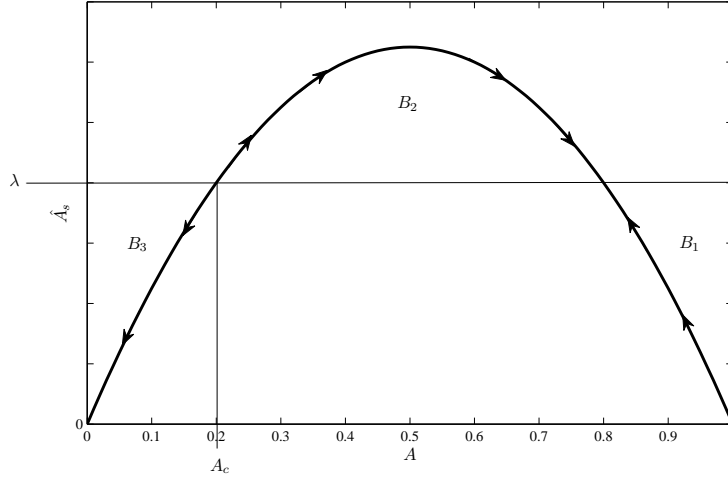


Figure 1: Areas of convergence and divergence of technology growth rates in the North and the South dependent on the distance to frontier

1 in area B_2 above the critical value A_c . A_s is larger than λ , so that the technology gap narrows, i.e. A increases. Economies fulfilling such initial conditions are able to catch up in terms of technologies up to a certain ratio $A < 1$, which fits to the growth performances of the "Asian Tiger" countries.

Starting on the left hand side of Figure 1 in area B_3 we identify the case of a growth trap. A is smaller than the critical value of A_c . We find the situation $\lambda > \hat{A}_s$ which leads to a movement on the parabola to the lower left hand side. The economy is too far away from the technology frontier. Moreover, the marginal product of capital in the South is too low relative to the marginal product of capital in the North to attract more foreign capital. This means that the developing economy is scarce in foreign capital and the technology diffusion speed is low due to an insufficient absorptive capacity. As a consequence, the South will end up with almost no foreign capital and very low technology diffusion. We observe this kind of behavior for example in the case of sub-Saharan African countries.

How can economic policy remedy such a convergence failure? One possibility is to increase the absorptive capacity H_s , which shifts the parabola in Figure 1 upwards so that the economy can move from a point in the divergence area to a point in the convergence area. A higher absorptive capacity prevents the convergence failure even in the long-run. The farther the South is away from the technology frontier, the more human capital is necessary to enable catching up.

Another possibility is to introduce own innovation with a certain rate in the follower country, which also shifts the parabola upwards. Hence, technology diffusion and

innovation positively interact as complements. But recalling equation (12), we notice that the last term steadily decreases over time as long as γ , the innovation rate in the South, is smaller than λ , the innovation rate in the North. λ is also the growth rate of A_s in the case of convergence. γ is likely smaller than λ , since the South is probably not as innovative as the North. Thus, at a certain point of time the South will fall back into the divergence area, when the difference between the technology level in the North and the technology level in the South has become too large. Alternatively, a subsidy on foreign capital gives an additional payoff to the foreign investment, i.e. is added to the marginal product of foreign capital expressed in (16). (For a discussion on promoting FDI see Hanson 2001.) The subsidy will be an effective remedy if it is high enough to overcome the critical point of a convergence failure. But again, only in the short-run. We can see from equation (16) that a constant subsidy becomes relatively unimportant when A_i and hence the marginal products of K_i grows in an exponential way. Therefore, the subsidy would have to rise together with technological progress in order too have a medium- or long-run effect.

5 Discussion

Our considerations along the line of the N&P theory describe that a better absorptive capacity as well as a higher volume of foreign capital narrow the technology gap between the North and the South until the technology growth rate of the South equals the growth rate of the technology frontier given by the North. In this case, there will be no reallocation of internationally mobile capital in the long-run. If the absorptive capacity of the South is below a certain threshold value, if the South is far away from the technology frontier and if certain preconditions are fulfilled, convergence of technology growth rates fails. As a consequence, the South falls further behind in terms of technologies, and foreign capital tends to be completely withdrawn from the South. The South is trapped.

Different to the literature so far (for example described by Benhabib and Spiegel 2005), we derive this outcome theoretically through the introduction of international capital mobility. A main contribution of our paper is therefore to reconcile the assumption that technological catching up is stronger the farther the distance to frontier with the alternative view that the diffusion speed is strongest at a medium distance to frontier (both discussed by Benhabib and Spiegel 2005) by introducing international capital

mobility. As a consequence, we identify the following trade-off: Far away from the technology frontier, there is a high potential for adopting new technologies, but there is also a lack of foreign capital. Close to the technology frontier, more foreign capital is allocated to the South, but there are fewer technologies left that can be adopted. Hence, the optimal technology diffusion speed is achieved somewhere at a medium distance to frontier. Therefore, our model is able to provide one possible explanation (among other explanations for example given by Hanson 2001) for the mixed findings in the empirical literature on technology spillovers via FDI and the empirical facts, for instance reported by the World Bank (2008): Some developing countries are able to catch up in terms of technologies, others are not, or even fall further behind.

We also examine the role of innovation in the South and its potential to narrow the North-South technology gap. In contrast to N&P, a higher initial level of technological progress in the North, interpreted as a better innovative capability of the North, widens the international technology gap in the presence of own innovation in the South. The reason is that own innovation in the South reduces the North-South technology gap which in turn reduces the possibility to exploit Northern technologies. In this sense, technology transfer and own innovation in the South are basically substitutes. They become complements in the short-run, since own innovations increase the marginal product of capital, which in turn attracts more foreign capital embodying advanced technologies. But in the long-run Southern innovation cannot prevent the South falling behind in terms of technologies, except when the South becomes as innovative as the North. This result is in accordance with Acemoglu et al. (2003a), stating that technological leaders follow an innovation-based strategy, while technological followers do not. Our long term outcome, that innovation is not an appropriate option for technological catching up of developing countries that lack in human capital, infrastructure and so forth, is in line with Acemoglu et al. (2003b). They show that imitation activities are more important far away from the technology frontier.

Our considerations also show that a larger income share of internationally mobile foreign capital is more beneficial for the South in terms of technology diffusion the smaller the North-South technology gap. Indeed, in China the revenue share of enterprises with Hong Kong, Macao, Taiwan and foreign funds in the revenue of all enterprises rose from 0.20 in 1998 to 0.24 in 2006 (China Statistical Yearbook 2006). Therefore, according to our stylized model, the Chinese policy of relaxing the requirements for FDI in China and supporting FDI inflows has been beneficial for technology diffusion. On the other hand, a constant subsidy on FDI inflows has only a short-run effect and

cannot prevent a convergence failure in the medium-run. The subsidy would have to rise in line with technological progress in order to stay effective. Taking the related fiscal burden into account, this sheds a somewhat critical light on such a policy in line with Hanson (2001). Thus, our outcomes support the need for a critical assessment of potential benefits of subsidizing FDI while considering the specific state of the economy in terms of existing technologies, human capital, infrastructure, the legal system and other country specific circumstances like the spatial distribution of economic activities in the light of agglomeration theory. Supporting FDI inflows in general without such a critical assessment can lead to substantial welfare losses for recipient countries.

Under the assumption of an economy that is closed for FDI inflows the model collapses to the standard Solow (or alternatively Ramsey) model where growth is driven by (exogenous) local innovation and domestic capital accumulation. In this sense the model is in line with success stories of economic growth without allowing significant foreign capital inflows such as in Japan and Korea.

However, the analysis is based on a simple stylized macroeconomic model that neglects endogenous innovation in the North and other channels of technology diffusion besides international capital mobility. It cannot capture other determinants of capital mobility (or FDI) besides returns on capital, either. Furthermore, capital transfer, absorption of technologies and innovation are costly. And keeping the absorptive capacity on a sufficiently high level requires permanent investment. In order to set up the optimal policy mix, policy makers need to know such costs, which are certainly hard to quantify. However, the previous analysis is not a cost benefit analysis, but rather a qualitative policy assessment.

Moreover, the analysis focusses solely on positive productivity spillover effects through *technology diffusion*. As pointed out by Hanson (2001), the entrance of foreign (multinational) enterprises can have negative effects on domestic enterprises and may lower the welfare of the host country because of several reasons. If the production factors intensively used by foreign enterprises are in inelastic supply, the entrance of foreign enterprises will put upward pressure on the related factor prices and thus increase production costs for domestic and foreign enterprises. Furthermore, in case of horizontal FDI, foreign enterprises may occupy a higher market share from domestic enterprises if the demand for their final products is in inelastic supply. Finally, the entrance of multinational firms likely increases competition in a non-competitive local market. The increased competition on the one hand enhances efficiency, on the other hand diminishes profits of domestic enterprises. Since multinational firms repatriate their profits, host

country welfare may decline. Such aspects cast doubt on a purely optimistic point of view of FDI that results in policies of promoting FDI inflows in general. However, such aspects cannot be captured by our stylized model.

Empirically, vertical productivity spillovers within the production chain seem to be more significant than horizontal spillovers between competing firms within sectors (for instance Javorcik 2004). In the case of vertical spillovers, negative impacts of FDI on rivaling domestic firms are less likely since these firms operate within the same production chain and not as rivals. Therefore, the distinction between horizontal and vertical linkages is important for economic policy. Furthermore, FDI seeking for the exploration of natural resources has a low potential for creating positive technology spillovers. However, our stylized macroeconomic model is not able to disentangle different kinds of FDI and spillovers within host economies.

Our considerations recommend opening developing countries for FDI inflows and improving the absorptive capacities at the same time in order to benefit from international technology diffusion successfully and to avoid poverty traps. The question remains why this does not happen in reality in many cases, for instance in Sub-Saharan African countries. Some obstacles for international transfer of capital and technologies are naturally given, such as landlockedness. Other factors such as education and infrastructure could be improved, but they are not because governments do not take the responsibilities. But this argument is not directly valid for opening the economy for FDI, because there are no directly related costs like in the case of educational investment. Nevertheless, FDI may create negative effects for host economies as described above. Consequently, certain stakeholders may be worse off due to the entrance of foreign enterprises and therefore lobby against opening the economy for FDI inflows, although such a policy might be beneficial for the economy as a whole (compare Das 1987).

6 Conclusion

We have analyzed a stylized macroeconomic model of North-South technology diffusion via capital mobility. The results show that one cannot rely on market forces as a guarantee for convergence of the Southern technology growth rate with the Northern technology growth rate via international capital mobility. Convergence of growth rates requires a sufficiently high absorptive capacity of the South (via education, infrastructure, legal framework etc.) in order to adopt new technologies successfully (additional

to certain preconditions discussed in the paper). Hence, according to the model it is not sufficient to rely on market driven international capital mobility or just to transfer capital and embodied technologies to developing countries. Such policies can be a waste of resources, if development policy does not ensure that the absorptive capacities of the recipient countries determined by education, infrastructure, the legal framework and other factors suffice.

Furthermore, fostering own innovation in the South positively interacts with technology diffusion in the short-run. But it cannot prevent the South falling behind in terms of technologies in the long-run, except when the South becomes as innovative as the North. This outcome confirms the importance of improving technology diffusion to developing countries rather than trying to create own innovations within developing countries, at least in early stages of development. Similarly, a subsidy on internationally mobile foreign capital yields a positive short-term effect that vanishes in the long-run. In order to stay effective, the subsidy would have to rise together with technological progress over time.

Moreover, the analysis suggests that a larger income share (possibly up to about 0.5) devoted to foreign capital is more beneficial when the technology level is closer to the technology frontier. This means that financial and tax advantages for foreign investors can be ill-designed, when the developing economy lacks in existing basic technologies and in absorptive capacity which are both necessary to attract foreign investment.

Further research could analyze the diffusion mechanism studied in this paper in an inter-temporal optimization framework including costs of capital transfer and innovation. It would yield the optimal allocation of foreign capital to the South along the optimal time path. The distinction of vertical and horizontal FDI as well as vertical and horizontal spillovers seems to be a fruitful avenue for further research. One could also attempt to include the technology diffusion framework of this paper into a general equilibrium framework with factor and product markets. This would allow to contrast the positive external effect of technology spillovers with possible negative effects of FDI for recipient countries. Moreover, a Computable General Equilibrium (CGE) model analysis could apply the diffusion mechanism to real data and reveal country and sector specific differences in the technology diffusion behavior.

7 References

- Acemoglu, D., P. Aghion and F. Zilibotti, (2003a). Distance to Frontier, Selection, and Economic Growth. *Journal of the European Economic Association* 4(1), 37-74.
- Acemoglu, D., P. Aghion and F. Zilibotti (2003b). Vertical Integration and Distance to Frontier. *Journal of the European Economic Association* 1(2-3), 630-638.
- Aghion, P. and P. Howitt (2005). Growth with Quality-Improving Innovations: An Integrated Framework. In: P. Aghion and S. Durlauf, eds., *Handbook of Economic Growth*, Elsevier, chapter 2.
- Barro, R. and X. Sala-i-Martin (1997). Technological Diffusion, Convergence, and Growth. *Journal of Economic Growth* 2, 127.
- Barro, R. J. and X. Sala-i-Martin (2004). *Economic Growth*. 2nd edition, MIT Press, Cambridge, Mass.
- Benhabib, J. and M. Spiegel (1994). The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data. *Journal of Monetary Economics* 34, 143-173.
- Benhabib, J. and M. Spiegel (2005). Human capital and technology diffusion. In: P. Aghion and S. Durlauf, eds., *Handbook of Economic Growth*, Elsevier, chapter 13.
- Blomström, M. and A. Kokko (1998). Multinational Corporations and Spillovers. *Journal of Economic Surveys* 12(3), 247-277.
- Borensztein E., J. De Gregorio and J.W. Lee (1998). How does foreign direct investment affect economic growth? *Journal of International Economics* 45(1), 115-135.
- Chen, E. K. Y. and J. H. Dunning, eds. (1994). *Technology Transfer to Developing Countries*. United Nations Conference on Trade and Development Programme on Transnational Corporations. Taylor and Francis.
- China Statistical Yearbook (2006). China Statistics Press.
- Ciruelos A. and M. Wang (2005). International Technology Diffusion: Effects of Trade and FDI. *Atlantic Economic Journal*, International Atlantic Economic Society 33(4).
- Crespo J., C. Martin and F. Velazquez, (2004). International technology spillovers from trade: the importance of the technological gap. *Investigaciones Economicas*, Fundacion SEPI 28(3), 515-533.
- Crispolti, V. and D. Marconi (2005). Technology transfer and economic growth in developing countries: an econometric analysis. Banca d'Italia, Temi di discussione series No. 564.
- Das, S. (1987). Externalities, and Technology Transfer through Multinational Corporations: A Theoretical Analysis. *Journal of International Economics* 22, 171-182.

- Diao, X., J. Ratts and H. E. Stokke (2005). International spillovers, productivity growth and openness in Thailand: an intertemporal general equilibrium analysis. *Journal of Development Economics* 76, 429-450.
- Findlay, R. (1978). Relative Backwardness, Direct Foreign Investment, and the Transfer of Technology: A Simple Dynamic Model. *The Quarterly Journal of Economics* 92(1) 1-16.
- Gerschenkron, A. (1962). *Economic Backwardness in Historical Perspective*. Harvard University Press.
- Girma, S., D. Greenaway and K. Wakelin (2001): Who benefits from Foreign Direct Investment in the UK? *Scottish Journal of Political Economy*, 48, 119-133.
- Girma, S. (2005). Absorptive capacity and productivity spillovers From FDI: a threshold regression analysis, *Oxford Bulletin of Economics and Statistics* 67(3), 281-306.
- Girma, S. and H. Görg (2007). The role of the efficiency gap for spillovers from FDI: Evidence from quantile regressions. *Open Economies Review* 18(2), 215-232.
- Grossman, G. and E. Helpman (1991). *Innovation and Growth in the Global Economy*. The MIT Press.
- Griffith, R., S. Redding and H. Simpson (2002). Productivity convergence and foreign ownership at the establishment level. IFS Working Paper WP02/22.
- Griffith R., S. Redding and J. van Reenen (2004). Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries. *The Review of Economics and Statistics*, MIT Press 86(4), 883-895.
- Hanson, G. H. (2001). Should Countries Promote Foreign Direct Investment? United Nations Conference on Trade and Development, Center for International Development, Harvard University, G-24 Discussion Paper Series.
- Hoekman, B. and B. S. Javorcik, eds. (2006). *Global Integration & Technology Transfer*. The World Bank, Washington DC.
- Javorcik, B. S. (2004). Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages. *American Economic Review* 94(3), 605-627.
- Keller, W. (2004). International Technology Diffusion. *Journal of Economic Literature*, American Economic Association 42(3), 752-782.
- Kneller, M. (2005). Frontier technology, Absorptive Capacity and Distance, *Oxford Bulletin of Economics and Statistics* 67(1), 1-23.
- Kokko, A. (1992). Foreign Direct Investment, Host Country Characteristics, and Spillovers. *Stockholm School of Economics*.
- Kokko A., R. Tansini and M. C. Zejan (1996). Local technological capability and productivity spillovers from FDI in the Uruguayan manufacturing sector. *Journal of*

Development Studies 32(4), 602-611.

Krugman, P. (1979). A Model of Innovation, Technology Transfer, and the World Distribution of Income. *Journal of Political Economy* 87(2), 253-266.

Lai, M., S. Peng and Q. Bao (2006). Technology spillovers, absorptive capacity and economic growth. *China Economic Review* 17, 300-320.

Lall, S. and S. Urata (2003). Competitiveness, FDI and Technological Activity in East Asia. Edward Elgar, Cheltenham, Northampton.

Mayer-Foulkes, D. and P. Nunnenkamp (2009). Do Multinational Enterprises Contribute to Convergence or Divergence? A Disaggregated Analysis of US FDI. Forthcoming in: *Review of Development Economics* 13(2), 304-318.

Nelson, R. and E. Phelps (1966). Investment in Humans, Technological Diffusion, and Economic Growth. *The American Economic Review, Papers and Proceedings*, 61, 69-75.

OECD (2002). Foreign Direct Investment for Development - Maximising Benefits, Minimising Costs. OECD Publications, Paris.

Romer, P. (1990). Endogenous Technological Change. *Journal of Political Economy* 98, 71-102.

Saggi, K. (2002). Trade, Foreign Direct Investment, and International Technology Transfer: A Survey. *World Bank Research Observer*, Oxford University Press 17(2), 191-235.

Van Meijl, H. and F. van Tongeren (1999). Endogenous International Technology Spillovers And Biased Technical Change In The GTAP Model. GTAP Technical Paper No. 15.

Xu, B., J. Wang J. (2000). Trade, FDI, and International Technology Diffusion. *Journal of Economic Integration* 15(4), 585-601.

Wang, J.-Y. and M. Blomström (1992). Foreign investment and technological transfer: a simple model, *European Economic Review* 36, 137-155.

World Bank (1993). The East Asian Miracle: Economic Growth and Public Policy. World Bank Policy Research Report. Washington DC.

World Bank (2008). Global Economic Perspectives. Technology Diffusion in the Developing World. Washington DC.

8 Appendix

In the following we examine the accumulation of domestic, internationally immobile, domestic (low-tech) capital D_i referring to the model presented in section 4.1. We

assume a constant savings rate s a la Solow and a depreciation rate δ .¹⁶

$$\dot{D}_i = sY_i - \delta D_i \quad (23)$$

The division of this equation by $\tilde{A}_i H_i$ yields a reformulation in efficiency units $k_i = \frac{K_i}{\tilde{A}_i H_i}$ and $d_i = \frac{D_i}{\tilde{A}_i H_i}$:

$$\frac{\dot{D}_i}{\tilde{A}_i H_i} = s(k_i^\alpha + d_i^\alpha) - \delta d_i$$

We transfer this expression following Barrow and Sala-i-Martin (2004):

$$\begin{aligned} \dot{d}_i &= \frac{\dot{D}_i}{\tilde{A}_i H_i} - \hat{A}_i d_i \\ \Rightarrow \dot{d}_i &= s(k_i^\alpha + d_i^\alpha) - (\hat{A}_i + \delta)d_i \end{aligned}$$

In the long-run, we find $\lim_{t \rightarrow \infty} (k_i^\alpha + d_i^\alpha) = d_i^\alpha$ since D_i is accumulated, while K_i is not. Thus, setting $\dot{d}_i = 0$ implicitly yields the following volume of domestic capital in efficiency units in the asymptotic steady state:

$$s(d_i^*)^\alpha = (\hat{A}_i + \delta)d_i^* \quad (24)$$

Accordingly, D_i asymptotically grows with the same rate as A_i , and Y_i grows with the same rate, too. In case of convergence, the Southern and the Northern economy therefore grow with the same rate A_n which is determined by the technological progress in the North. The steady state value d_i^* rises with the propensity to save and invest and declines with the depreciation rate and with the rate of technological progress. As a consequence, the ratio of domestic capital to the technology level declines when the rate of technological progress becomes higher.

Moreover, we are able to determine the volume of high-tech capital K_s in absolute terms (not only for the steady state but for any point of time) using equation (18) and our initial assumption $K_n + K_s = 1$:

$$\begin{aligned} \frac{K_s}{1 - K_s} &= \tilde{A}H \\ \Leftrightarrow K_s &= \frac{\tilde{A}H}{1 + \tilde{A}H} \end{aligned} \quad (25)$$

¹⁶Alternatively, capital could be accumulated via an inter-temporally optimal choice of consumption a la Ramsey.

The closer the Southern technology level catches up to the Northern level and the higher the absorptive capacity of the South, the more high-tech capital is obviously allocated to the South.

Finally, Figure 2 plots \hat{A}_s as a function of A for different values of α .

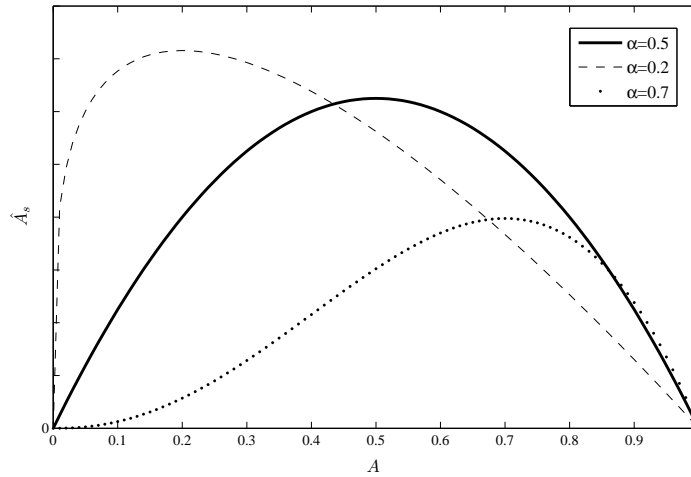


Figure 2: Rate of technological progress in the South dependent on the distance to frontier for different income shares of foreign capital in the South

Chapter 3:
Energy Savings via FDI?
Empirical Evidence from Developing Countries

Published in Environment and Development Economics, Cambridge Journals, 15(1),
59-80.

<http://journals.cambridge.org/action/displayIssue?jid=EDE&volumeId=15&issueId=01&iid=7150020>

Chapter 3:

Energy Savings via FDI?

Empirical Evidence from Developing Countries ^{*}

Michael Hübler [†] & Andreas Keller [‡]

August 25, 2009

Abstract

In this paper we examine the influence of foreign direct investment inflows on energy intensities of developing countries empirically. We first replicate a simple OLS estimation, as it is found in the literature, that suggests energy intensity reductions from FDI inflows. However, the OLS estimation turns out to be spurious and only a starting point for further research. In our regressions we use macro level panel data on 60 developing countries for the period 1975-2004 including other potential determinants of energy intensities and carry out robustness checks with more specific data. The results do not confirm the hypothesis that aggregate FDI inflows reduce energy intensities of developing countries. Rather, foreign development aid seems to be related to energy efficiency gains.

JEL Classifications: F18, F21, O13, O33, Q43

Keywords: Developing countries, energy intensity, FDI, technology transfer

* We thank Sonja Peterson, Johannes Bröcker, Christoph Böhringer, Andreas Waldkirch, Peter Nunnenkamp, Ulrich Oberndorfer, Kai Carstensen, Erich Gundlach, Gernot Klepper, Jonas Dovern, Toman Omar Mahmoud and two anonymous reviewers for helpful comments and suggestions. We thank Stefan Bickert for his data support.

[†] Kiel Institute for the World Economy, Düsternbrooker Weg 120, 24100 Kiel, Germany,
E-mail: michael.huebler@ifw-kiel.de, Tel: +49-431-8814-401.

[‡] University of Oldenburg, Department für Wirtschafts- und Rechtswissenschaften, 26111 Oldenburg, Germany,
E-mail: andreas.keller@uni-oldenburg.de, Tel: +49-441-798-4445.

1 Introduction

While economic growth in developing countries is a desirable goal, its side effects of rising energy demand and greenhouse gas emissions are problematic in light of global warming. The controversial discussions during the Bali climate policy conference in 2007 showed that including the developing countries in a post-Kyoto agreement in a fair way is a challenge, and it was emphasized that the industrialized countries need to support the developing countries by technical, financial and educational measures.

One possibility to slow down rising emissions is energy saving technology transfer to developing countries since energy use is strongly related to carbon emissions. It is thus important to detect how international technology transfer occurs and how it affects energy supply and demand. Foreign direct investment (FDI) is regarded as one important channel for technology transfer in general (e.g. Keller, 2004). Empirical evidence on energy saving international technology transfer could answer the question whether more FDI flows to developing countries can effectively help to restrain energy use and thus greenhouse gas emissions. In this paper we use a large macro level panel data set and investigate empirically whether FDI inflows have a significant effect on the energy intensity of developing economies.

Despite the relevance of the topic, only little empirical evidence exists regarding the transfer of energy saving technologies. Some case studies and micro-econometric work using firm-level data indicate that foreign owned firms in specific developing countries indeed use less energy than their indigenous counterparts (Eskeland and Harrison, 2003; Fisher-Vanden et al., 2004). However, comprehensive studies on an aggregate level with cross-country data are missing. This is unfortunate since they can potentially give a broader picture of the issue at hand and can answer the question whether the findings from specific plant-level studies can be generalized on a global scale. The question of generalization is important in light of recent worldwide initiatives to reduce energy use and emissions. Mielnik and Goldemberg (2002) use data on 20 developing countries and find a negative effect of FDI on energy intensity. However, this analysis controls for no other influences on energy demand and does not take the unit root behavior of the variables into account. Hence, rather than providing general evidence, the results by Mielnik and Goldemberg open a promising way for further research.

In this paper we use a large macro level panel data set and an extended econometric model in order to broaden the analysis and to investigate whether FDI has a general impact on energy intensity. The data set includes 60 developing countries for the years 1975 to 2004. In contrast to the hypothesis proposed by Mielnik and Goldemberg (2002), we find no robust energy reducing effect of FDI inflows in developing countries. The interactions of FDI inflows with country-specific characteristics yield no significant results, either. For policy makers who seek to achieve energy reductions in developing countries, our results imply that a general support of FDI inflows is not enough; it is rather necessary to explicitly encourage that kind of foreign direct investment that brings about energy reducing technology transfer. Issue-linkage, as for example the clean development mechanism in the Kyoto Protocol, might be a way to achieve this objective.

Our paper is structured as follows. Section 2 explains the theoretical background and reviews related empirical evidence in the literature. We furthermore replicate a simplified regression analysis as in Mielnik and Goldemberg (2002) and show why an extended empirical approach is needed to generate meaningful results. In section 3, we first derive our empirical model based on the insights from the theoretical considerations. We then specify our estimating equations, apply regression analyses and discuss the results. Section 4 presents our robustness checks using a modified database. Section 5 concludes.

2 The influence of FDI inflows on energy use

In this section we first present the conceptual background for the hypothesis that foreign direct investment has an energy reducing effect in the destination country. Then we give an overview of the empirical literature examining the evidence on energy reducing technology transfer via FDI and discuss in more detail the analysis by Mielnik and Goldemberg (2002).

2.1 Theoretical background

In the literature on trade and the environment, it has become common to decompose the effects of economic activity and trade on pollution into a scale, a composition and a technique effect. This decomposition was introduced by Grossman and Krueger (1993), who examined data on concentrations of sulphur dioxide and suspended particulate matter. It was later on backed up with formal theory by Antweiler et al. (2001) and further elaborated by Copeland and Taylor (2003). As argued by Cole (2006), the framework from the trade and environment literature is also applicable to energy use since energy use is the principal cause of most air pollutants. In our analysis, we adapt the concept developed by Antweiler et al. (2001) and examine how scale, composition and technique effects influence total energy use.

For simplicity, we split up total economic activity into two sectors with different energy intensities: industry and non-industry. Energy use in the non-industry sector relative to its output value is defined by $e(A)$ with the properties $e(A) > 0$ and $e'(A) < 0$, where A is a proxy for the average technology in use. We assume that the energy intensity in the industry sector is always μ times higher than $e(A)$. Measuring total output by GDP and denoting the share of industrial value added in output by IND , total energy use can be written as:

$$E = GDP \cdot (\mu \cdot IND + 1 - IND) \cdot e(A) \quad (1)$$

It follows from (1) that total energy use in a country can be decomposed into three factors: the scale of overall economic activity (GDP), the relative importance of energy-intensive sectors in economic activity (IND), and the energy intensity of the technology in use ($e(A)$). In analogy to Antweiler et al. (2001), changes in these factors can be interpreted in the following way: The first term is obviously the scale effect and reflects the impact of increasing or decreasing economic activity on total energy use when holding constant the mix of sectors as well as the technology. The second term is the composition effect: Holding scale and technology constant and producing relatively more energy-intensive goods due to a sectoral shift leads to an increase in energy use. The last term captures the technique effect resulting from the implementation of a more energy efficient technology A that reduces the input of energy into the production process.

Having set out the fundamental factors affecting energy use, we can now identify the variables that influence these factors. In this context, we pay special attention to the effects of foreign direct investment inflows.

The *scale effect* can be directly measured by changes in the observable variable *GDP*. For constant returns to scale of production functions, an output increase leads to a proportional increase in energy inputs. This means that energy intensity, defined as energy use divided by GDP, remains constant.¹ Foreign direct investment can influence the scale effect indirectly: Since FDI inflows are assumed to stimulate economic growth and since expanded economic activity is related to higher energy use, the scale effect resulting from FDI inflows is positive and its extent depends on the magnitude of the influence of foreign direct investment on GDP.

Changes in the relative importance of the production of energy-intensive goods, i.e. the *composition effect*, are determined by changes in the sectoral structure of an economy. A typical empirical observation is that in early stages of a country's economic development, economic activity shifts from the agricultural to the industrial sector. Since the latter is more energy intensive than the former, this implies a positive composition effect. Later in the development process, activity moves typically from the industry to the service sector or from the heavy to the lighter industry. This implies a negative composition effect in this stage as the service sector and the light industry are less energy intensive (Stern, 2004). One reason for such a development pattern can be changes in the comparative advantage of an open economy in the world market. The comparative advantage can be influenced by variables like the capital-to-labor ratio, environmental regulations, or the share of skilled labor. Foreign direct investment inflows contribute to the composition effect if they change the sectoral structure of an economy. In equation (1) we measure sectoral shifts directly by changes in the share of industrial value added in GDP. This means that all the indirect influences on the sectoral structure, including the composition effect of FDI, are already captured by the inclusion of this variable.

The *technique effect* covers the impact of employing new technologies or management practices on energy use. We hypothesize that novel technologies developed with a higher level of knowledge are more productive and hence more energy efficient than old technologies. In analogy to Arrow (1962), we propose that

learning is a function of cumulated gross investment, which we denote by G_I .² This means that accumulated knowledge increases with every new capital good being available. The hypothesis brought forward by Mielnik and Goldemberg (2002) is that imported investments via FDI in developing countries have a stronger energy reducing effect than new domestic investments. This can be the case through *technology transfer* via FDI so that developing countries are able to ‘leapfrog’ over traditional technologies in use. In order to test this hypothesis, we explicitly consider the accumulated FDI inflows G_{FDI} as a determinant of energy intensity, which represents a source of accumulated knowledge that might differ from accumulated domestic investment in its strength. We furthermore add the accumulated import value G_{IM} and the accumulated foreign aid value G_{AID} as further potential sources of technology transfer. All three variables are expressed as shares relative to overall cumulative gross investment G_I since the shares of foreign capital (embodied in FDI, imports or aid) in total capital represent the relative importance of transferred knowledge in the economy (similar Aitken and Harrison, 1999). A different driver of the energy saving technique effect can be income per capita, which can be explained by a political economy mechanism: Since environmental quality is assumed to be a normal good, rising per-capita income leads to higher public appreciation of a clean environment. The rising demand for environmental quality may then result in the adoption of stricter environmental regulation that typically also provides incentives to reduce energy use.³ Firms react to the regulation by the introduction of cleaner and more energy efficient technologies. This effect, which is related to the so-called environmental Kuznets curve, is in the following referred to as *income induced technique effect*.⁴ Summing up, we expect that the energy intensity of a country’s technology $e(A)$ depends negatively on accumulated gross investment G_I and might

¹ Note that this implication would not hold in the case of economies of scale, i.e. when an expansion of output needs a proportionally lower or higher increase in input quantities.

² Other variables that are typically used as a proxy for learning and technical progress are accumulated GDP or GDP per capita.

³ This is obvious in the case of energy taxes but applies as well to other environmental policy measures like stricter regulations of pollutant emissions arising as a by-product from the use of fossil energy resources.

⁴ The environmental Kuznets curve (EKC) stands for the impact of a country’s per capita income on (per capita) pollution suggesting an inverted U-shaped relationship. Although the EKC is a well-known concept and is regarded as a stylized fact in environmental economics, its existence has recently been challenged on both theoretical and empirical sides (e.g. Stern, 2004; Siebert, 2005). The environmental Kuznets curve has traditionally been applied to emissions of local pollutants, but recent studies also apply this concept to CO₂ emissions (e.g. Mazzanti et al., 2006) as well as energy *intensity* (Galli, 1998).

be reduced further by the extent of the respective shares of accumulated FDI, imports and aid in accumulated investment. Furthermore, the income level per capita (YPC) might reduce energy intensity so that $e(A)$ is a function of:

$$e(A) = f \left(G_I, \frac{G_{FDI}}{G_I}, \frac{G_{IM}}{G_I}, \frac{G_{AID}}{G_I}, YPC \right) \quad (2)$$

With regard to FDI, it is also possible to establish an indirect link to the income induced technique effect, namely under the assumption that openness to trade and foreign direct investment enhances growth and hence per capita income. Grossman and Krueger (1993) emphasize especially this indirect effect when they analyze the technique effect of trade liberalization. In the theoretical model by Antweiler et al. (2001), the income induced technique effect is the *only* technique effect resulting from increasing openness.⁵ While the literature on trade and the environment has mainly focused on the income induced technique effect, the broader literature on technological change has recently expanded with numerous studies on technology transfer resulting from openness to trade and foreign direct investment (for an overview see Keller, 2004). Technology transfer via FDI, which potentially also reduces energy use, can occur in two ways: First, directly via more efficient foreign firms operating in the host country and second, indirectly through technological spillovers from foreign firms to indigenous firms. For these mechanisms to be effective, we make the assumption that the technology used by foreign investors is typically superior to the technology that is currently in place in developing countries. The direct effect implies that the foreign owned firm, compared to a similar indigenous firm, uses less energy and hence contributes to the technique effect. Regarding the technological spillovers, the literature suggests three potential channels: demonstration effects, which stand for imitation and reverse engineering by local firms, labor turnover, which implies the transfer of knowledge by workers who change their employer, and vertical linkages, which involve that multinationals transfer technology to their suppliers or customers (Saggi, 2002). Additionally, higher exports and imports as well as foreign direct investment inflows likely lead to

⁵ In the empirical estimation of their model, Antweiler et al. (2001) provide a sensitivity test where they explicitly allow for a direct effect of FDI on SO_2 concentrations. However, they find no substantial relationship between the extent of FDI in an economy and the pollution level.

increased competition. Firms need to become more productive in order to stay competitive in the export market or to compete with imports and new foreign companies in the domestic market. For instance, Corcos et al. (2007) apply a theoretical model with heterogeneous firms and assume that international trade increases aggregate productivity through a selection effect: The least productive firms leave the market under increased pressure from competition. One can expect that higher productivity also implies more efficient energy use in production.

Though we identified several indirect influences of FDI inflows on energy use through scale, composition and technique effects, note that all the indirect effects are implicitly included in the variables *GDP* (scale effect), *IND* (composition effect) and *YPC* (income induced technique effect). We therefore propose that any remaining direct influence of FDI can be attributed to technology transfer and the main contribution of our analysis is to identify and quantify this influence empirically.

2.2 Empirical literature on energy saving technology transfer via FDI

There is a large and growing empirical strand of literature on productivity improving technology transfer and spillovers through FDI. The evidence is mixed, but Keller (2004) notes in his survey article that “recent micro productivity studies tend to estimate positive, and in some cases also economically large spillovers associated with FDI” (p. 771).⁶ Among other studies, Tybout (2002) finds evidence for efficiency improvements due to higher exposure to foreign competition as predicted by the New Trade Theory. However, he points out that it is not clear according to the related literature whether international activities cause the improved efficiency or vice versa. Furthermore, none of these empirical studies sheds light on the effects of spillovers and productivity gains on energy efficiency.

In the following, we focus specifically on *energy saving* technology transfer via FDI. Peterson (2008) reviews the existing evidence and remarks that although there may be a large potential for such technology transfer, there is a lack of knowledge about its empirical magnitude and its drivers, and the topic remains insufficiently researched. The hypothesis that foreign owned companies use less

energy than their indigenous counterparts in developing countries is confirmed by studies based on firm-level data. In their analysis of manufacturing plants in Cote d'Ivoire, Mexico and Venezuela, Eskeland and Harrison (2003) find that foreign ownership is associated with less energy use. A similar result is documented by Fisher-Vanden et al. (2004), who find a negative impact of foreign ownership on the energy intensity of Chinese companies. These examples suggest that the more efficient technologies of foreign firms can indeed contribute to an energy reducing technique effect via technology transfer.

On an aggregate level, only very few studies link openness and FDI to energy saving technology transfer. Cole (2006) uses a variation of the model developed by Antweiler et al. (2001) to examine the impact of trade intensity (while not explicitly including FDI) on energy use in 32 developed and developing countries. His panel estimation yields that the effect of liberalization is country-specific and can be positive or negative, depending on whether the country is importing or exporting the energy-intensive good.

Based on previous evidence that developing and industrialized countries are converging to a common pattern of energy use (Mielnik and Goldemberg, 2000), Mielnik and Goldemberg (2002) carry out a simplified regression as a starting point for further research. Their results indicate that the quantity of foreign direct investment inflows has a negative influence on energy intensity. They use a sample of 20 developing countries⁷ for the years 1987 to 1998, aggregating all countries to one time series, and estimate the regression:

$$EI_t = \beta_0 + \beta_1 \frac{FDI_t}{I_t} + \varepsilon_t \quad (3)$$

The dependent variable EI_t is energy intensity in year t , i.e. the sum of total energy use in all 20 countries divided by their GDP, which is measured in purchasing power parity. The explanatory variable FDI_t / I_t represents inflows of foreign direct investment as a fraction of total gross investment in all countries. ε_t is the error term.

⁶ Aitken and Harrison (1999), Javorcik (2004), Javorcik and Spatareanu (2008) and Keller and Yeaple (2009) are prominent examples for examinations of spillovers to domestic firms via FDI.

⁷ The countries are in alphabetical order: Algeria, Brazil, Chile, China (PR), Colombia, Costa Rica, Egypt, India, Indonesia, Malaysia, Mexico, Morocco, Nigeria, Pakistan, Peru, Philippines, Singapore, South Africa, Thailand, Uruguay.

First, we reproduce the regression by Mielnik and Goldemberg with the same 20 countries, but for an extended time span. Complete data on gross fixed capital formation (gross investment) is available for each of the 20 countries from 1979 to 2003. We evaluate our dataset with the same regression model as Mielnik and Goldemberg and find a similar result, which supports the view of a strong energy reducing impact of foreign direct investment: FDI_t / I_t has a negative coefficient β_1 of -0.774, the constant β_0 is 0.320, and R^2 is 0.818. However, the result of this OLS estimation relies critically on the stationarity assumption of the involved variables.⁸ If at least one variable is instead integrated, which implies non-stationarity, standard OLS regression analysis is not appropriate and can result in a spurious regression (compare Granger and Newbold, 1974).

We check the variables for the stationarity property via the Augmented Dickey-Fuller (ADF) test for unit roots. According to obvious trends in the data, we test both variables against stationarity around a deterministic trend.⁹ The null hypothesis of a unit root cannot be rejected in all tests of the variables in levels, but can be rejected in case of the first differences. Thus, the variables energy intensity and foreign direct investment as a fraction of total investment are both integrated of order one. Furthermore, we employ the Johansen cointegration test. Using the variables EI_t and FDI_t / I_t and examining three possible model specifications, no cointegration can be found in any case. The results of the preceding tests show that the estimation results obtained from the classical OLS regression are likely to be misleading and we need to treat a simplified regression analysis with caution.

3 Panel data analysis

In order to find empirical evidence for energy saving international technology transfer via foreign direct investment, we use aggregated country-level data for 60 developing countries in the period 1975 to 2004. Thereby we refer to all Non-Annex I Parties of

⁸ A time series $X = \{x_1, \dots, x_m\}$ is called stationary if it has

1. a constant and finite mean over time: $E[x_t] = \mu$ for all t
2. a constant and finite variance over time: $\text{Var}[x_t] = \sigma^2$ for all t
3. and constant covariances over time: $\text{Cov}[x_t, x_{t+s}] = \sigma_s^2$ for all t, s .

⁹ For the additional Figure A1 illustrating the trends of both variables and for the additional Tables A1 and A2 with the results of the test statistics please see the Online Appendix at <http://journals.cambridge.org/EDE>.

the Kyoto Protocol as developing countries. Excluding countries with insufficient data availability (among them all countries that have emerged from the former Soviet Union) results in our sample of 60 countries.¹⁰ Data on total primary energy supply are taken from the International Energy Agency (2007). Data on all other variables are found in the World Development Indicators by the World Bank (2007) and, if not available there for specific countries, from the Balance of Payments Statistics and the International Financial Statistics of the International Monetary Fund (2007a; 2007b). An exact definition of the variables and units used in the regressions is presented in Table 2 in the Appendix. For a detailed data description, please see the Online Appendix at <http://journals.cambridge.org/EDE>.

In our extended empirical specifications we take potential non-stationarity of the time series into account and use several control variables derived from our theoretical framework. Moreover, we account for the heterogeneity of countries by applying panel data methods. This is recommendable since the aggregation of all countries would eliminate heterogeneity, which is obviously present in our sample of 60 countries.

3.1 Derivation of the empirical model

We derive our empirical model for energy use from the theoretical framework and the variables presented in section 2.1. Hence, we need to add up the scale effect, the composition effect and the technique effect. In the case of constant returns to scale, the absolute size of an economy measured by GDP has no influence on energy intensity. We therefore leave the scale effect aside and analyze energy intensity EI as the dependent variable, i.e. total primary energy supply divided by GDP (as in Mielnik and Goldemberg, 2002; Cole, 2006), where GDP is measured in purchasing power parities.¹¹ Furthermore using a multiplicative form for equation (2), inserting it

¹⁰ The data are available upon request. The countries are in alphabetical order: Algeria, Angola, Argentina, Bangladesh, Benin, Bolivia, Botswana, Brazil, Cameroon, Chile, China (PR), Colombia, Congo (DR), Congo (Rep), Costa Rica, Cote d'Ivoire, Cyprus, Dominican Republic, Ecuador, Egypt, El Salvador, Gabon, Ghana, Guatemala, Honduras, India, Indonesia, Iran, Jamaica, Jordan, Kenya, Korea (Rep), Lebanon, Malaysia, Malta, Mexico, Morocco, Mozambique, Nicaragua, Nigeria, Oman, Pakistan, Panama, Paraguay, Peru, Philippines, Saudi Arabia, Senegal, South Africa, Sri Lanka, Syria, Tanzania, Thailand, Togo, Trinidad and Tobago, Tunisia, Uruguay, Venezuela, Zambia, Zimbabwe.

¹¹ For a discussion on purchasing power parities in measuring energy intensity see Birol and Okogu (1997).

in (1) and taking logs, the energy intensity of developing country i at time t can be expressed as:

$$\begin{aligned} \ln EI_{it} = & \ln(\mu \cdot IND_{it} + 1 - IND_{it}) + \alpha_1 \ln G_{I_{it}} + \\ & + \alpha_2 \ln\left(\frac{G_{FDI_{it}}}{G_{I_{it}}}\right) + \alpha_3 \ln\left(\frac{G_{IM_{it}}}{G_{I_{it}}}\right) + \alpha_4 \ln\left(\frac{G_{AID_{it}}}{G_{I_{it}}}\right) + \alpha_5 \ln YPC_{it} \end{aligned} \quad (4)$$

We are not able to include all various influences on energy use since many of them are determined on the micro level and the necessary micro data is not available for most developing countries. Examples are the detailed sectoral structure of the industry, local economic and environmental regulations or energy prices. The way how we will implicitly capture country- or time-specific effects in the estimation is to use panel data models with cross-section fixed or random effects as well as time fixed effects.

We test the available variables for the presence of a stochastic trend and thus on stationarity. We apply several unit root tests for panel data and allow for individual trends and intercepts.¹² Not surprisingly, we find unit roots for total primary energy supply, GDP and GDP per capita (as confirmed in other empirical studies, e.g. Perman and Stern, 2003).¹³ Regarding energy intensity (primary energy supply divided by GDP) the unit root hypothesis is rejected. However, the probability of error in this case is close to 5 percent, which indicates some uncertainties about the properties of this variable. As a consequence of the unit root tests' results, we transform our estimation model into first differences, where according to the test statistics all variables are stationary without doubt.

Approximating $\ln(\mu \cdot IND_{it} + 1 - IND_{it})$ in equation (4) by $(\mu - 1) \cdot IND_{it}$ and taking first time differences leads to:

¹² We employ unit root tests by Maddala and Wu (1999) and Choi (2001) (based on Dickey and Fuller, 1979; Phillips and Perron, 1988) as well as Levin, Lin and Chu (2001) and Im, Pesaran and Shin (2003).

¹³ Though total primary energy supply and GDP do not show up in equation (4), we will later estimate an alternative specification that explicitly includes the scale effect and therefore uses total primary energy supply as the dependent variable and GDP as one of the explanatory variables.

$$\begin{aligned} \Delta \ln EI_{it} = & (\mu - 1) \cdot \Delta IND_{it} + (\alpha_1 - \alpha_2 - \alpha_3 - \alpha_4) \frac{I_{it}}{G_{I_{it}}} + \\ & + \alpha_2 \frac{FDI_{it}}{G_{FDI_{it}}} + \alpha_3 \frac{IM_{it}}{G_{IM_{it}}} + \alpha_4 \frac{AID_{it}}{G_{AID_{it}}} + \alpha_5 \Delta \ln YPC_{it} \end{aligned} \quad (5)$$

Note that $\Delta \ln$ can also be expressed by relative changes; for instance $I_{it} / G_{I_{it}}$ means gross investment flow in year t over cumulative gross investment up to year t . We now assume for reasons of data availability that accumulated flows of investment, FDI, imports and aid are proportional to the GDP of the respective economy in year t , for example $G_{FDI_{it}} = \sigma_2 Y_{it}$. Equation (5) then reads:

$$\begin{aligned} \Delta \ln EI_{it} = & (\mu - 1) \cdot \Delta IND_{it} + \frac{(\alpha_1 - \alpha_2 - \alpha_3 - \alpha_4) I_{it}}{\sigma_1 Y_{it}} + \\ & + \frac{\alpha_2 FDI_{it}}{\sigma_2 Y_{it}} + \frac{\alpha_3 IM_{it}}{\sigma_3 Y_{it}} + \frac{\alpha_4 AID_{it}}{\sigma_4 Y_{it}} + \alpha_5 \Delta \ln YPC_{it} \end{aligned} \quad (6)$$

Writing the coefficients in a more comprehensive way and, for convenience, denoting the shares of investment, FDI, import and aid flows in GDP by I_{it} , FDI_{it} , IM_{it} and AID_{it} leads to our empirical model:

$$\begin{aligned} \Delta \ln EI_{it} = & \underbrace{\beta_1 \Delta IND_{it}}_{\text{Composition effect}} + \underbrace{\beta_2 I_{it}}_{\text{Vintage capital effect}} + \underbrace{\beta_3 FDI_{it} + \beta_4 IM_{it} + \beta_5 AID_{it}}_{\text{Technology transfer}} + \underbrace{\beta_6 \Delta \ln YPC_{it}}_{\text{Income induced technique effect}} \end{aligned} \quad (7)$$

Linking the model to our conceptual framework, the coefficients of FDI_{it} , IM_{it} and AID_{it} encompass international technology transfer, while the coefficient of I_{it} also includes a domestic vintage capital effect that may contribute to a reduction of energy intensity. The composition effect is captured by ΔIND_{it} , the change in the share of industrial value added in GDP. The income induced technique effect, which FDI may also bring about via rising incomes, is accounted for by the relative change in per

capita income $\Delta \ln YPC_{it}$. The scale effect is implicitly included since we examine energy use relative to the size of the economy represented by GDP.

Our hypothesis of energy savings via FDI implies that β_3 is negative. The interpretation is straightforward: FDI inflows in a certain year decrease energy intensity in relative terms, whereas ceteris paribus energy intensity stays constant if there is no FDI inflow. The higher the FDI inflow intensity, the higher is the relative reduction of energy intensity. This is consistent with the notion that foreign direct investment inflows continuously bring about technology transfer that can reduce energy intensity. The coefficients of imports and aid can be interpreted in the same way.

3.2 The estimating equations

Since there is considerable heterogeneity in the development of energy intensities among countries, we allow for unobserved effects in the estimation. To check the need for country-specific effects we carry out *LR*-tests (likelihood ratio tests) and *F*-tests on poolability. (Note that country-specific effects in the differenced equation imply country-specific time trends in the equation in levels.) The tests reject the null hypothesis that fixed effects are redundant in all cases. We then compute Hausman tests for choosing between fixed and random effects, which show that random effects are consistent only in model specification A following below. In principle, changes in the relative price of energy should be considered since they have an effect on the technology and the relative employment of input factors. Unfortunately, we were not able to obtain energy prices over the sample period for most of the developing countries. Nevertheless, it is possible to include worldwide changes in energy prices indirectly via time-specific fixed effects. These effects are furthermore able to capture any other time-specific influences that affect all countries in the sample in a similar way.

In our parsimonious model specification A we follow Mielnik and Goldemberg and explain energy intensity by the variable FDI_{it} / I_{it} , i.e. foreign direct investment relative to total investment:

$$A: \Delta \ln EI_{it} = \alpha + \chi_i + \theta_t + \beta_1 \frac{FDI_{it}}{I_{it}} + \varepsilon_{it} \quad (8)$$

α is the overall constant, χ_i are country-specific effects, and θ_t are period-specific effects. The ε_{it} are the error terms. Following the Hausman test result, we start with cross-section random effects in regression A1 (without period-specific effects, because we cannot use both cross-section and period random effects in our unbalanced panel). Since we would also like to capture time dependent effects and since the F -/ LR -tests reject the redundancy of time dependent effects, we turn to specification A2 with both country and time fixed effects. As in Mielnik and Goldemberg (2002), we do *not* include further control variables in these specifications.

This estimation, however, will suffer from an omitted variable bias if other determinants of energy intensity are partially correlated with FDI and not captured by the country- or time-specific effects. Therefore we turn to specification B that directly follows from equation (7) and additionally includes a constant plus country- and time-specific effects. For the exact definition of the variables see Table 2 in the Appendix.

$$\begin{aligned} \text{B: } \Delta \ln EI_{it} = & \alpha + \chi_i + \theta_t + \beta_1 \Delta IND_{it} + \beta_2 I_{it} + \beta_3 FDI_{it} \\ & + \beta_4 IM_{it} + \beta_5 AID_{it} + \beta_6 \Delta \ln YPC_{it} + \varepsilon_{it} \end{aligned} \quad (9)$$

Throughout specification B, we apply cross-section and time fixed effects since on the one hand, the F -test rejects the hypotheses of redundant cross-section as well as period fixed effects and on the other hand, the result of the Hausman test suggests that random effects are inconsistent. The pair wise correlations between the explanatory variables are all low, so that multicollinearity should not be a problem (for the correlation matrix see Table A4 in the Online Appendix at <http://journals.cambridge.org/EDE>). The highest correlation of FDI with another variable is found regarding imports with 0.35. To be on the safe side, we always complement each regression by first omitting imports and then omitting FDI. In these additional regressions, we find no significantly different results than the results we present for the full spectrum of variables.

It is furthermore noteworthy that due to the estimation in first differences, the regression is not able to capture those technological spillovers that occur only with a time lag after the FDI inflow has been recorded. Knowledge transferred from abroad likely diffuses further within the country with a time delay. We also notice that GDP

is part of the dependent variable and several explanatory variables. This is appropriate as long as the resulting intensity variables develop independently from GDP over time. If GDP fluctuates in the short run while energy supply, FDI, imports, aid, the investment share and the industry share all adjust sluggishly, we can possibly detect a resulting correlation between energy intensity and the regressors since they are all influenced by GDP fluctuations. Especially, if energy intensity is strongly affected by short term GDP fluctuations, the effect of income, measured by GDP per capita, on energy intensity might be caused by the design of the variables.¹⁴ In order to remedy the potential problems, we employ specification B2 where we replace the values of all explanatory variables by their one-period lagged counterparts (B1 is the variant without time lags). This means that FDI inflows affect energy intensity in the year after the actual inflow.¹⁵ In order to explicitly allow for a longer time lag in spillover effects, we furthermore employ specification B3, where we use moving averages of the past three years for the variables FDI, imports and aid inflows.

In the alternative specification C, we use total primary energy supply E as the dependent variable and replace income per capita (YPC) by total GDP (Y) as an explanatory variable. In this case, Y captures both the income induced technique effect and the scale effect without imposing the restriction of constant returns to scale.¹⁶ An advantage compared with model B is that GDP is no longer part of the dependent variable, therefore reducing distortions caused by short-term GDP fluctuations.

$$C: \Delta \ln E_{it} = \alpha + \chi_i + \theta_t + \beta_1 \Delta IND_{it} + \beta_2 I_{it} + \beta_3 FDI_{it} + \beta_4 IM_{it} + \beta_5 AID_{it} + \beta_6 \Delta \ln Y_{it} + \varepsilon_{it} \quad (10)$$

Nevertheless, there is still the possibility of an endogeneity bias since the change in energy use might itself affect GDP or even FDI. Therefore, we again employ a second

¹⁴ According to the Durbin-Wu-Hausman endogeneity test, we find that GDP per capita is endogenous. We therefore use a TSLS (two stage least squares) estimation as a robustness check for specification B1 and employ GDP per capita lagged for one period as an instrument for current GDP per capita. Testing for the presence of a weak instrument by running a reduced form regression finds no indication of a weak instrument. The Sargan test for overidentification does not suggest including GDP per capita lagged for two periods as an additional instrument.

¹⁵ Note that an inflow of FDI typically takes place at one specific point in time during the year. If the inflow is recorded at the end of the year it is reasonable that the effect on energy intensity takes place only in the following year. Furthermore, a lag of one year takes delayed spillovers at least partially into account.

specification C2 where we replace the values of all explanatory variables by their one-year lagged counterparts, while C1 is without time lags. To furthermore test for delayed spillover effects in technology transfer, we employ specification C3 with moving averages of the past three years for FDI, imports and aid. Note that specifications C2 and C3, in contrast to all other specifications, do not control for contemporary changes in real GDP so that the scale effect of FDI is implicitly included in the coefficient of FDI.¹⁷

3.3 Regression results

The results referring to specifications A, B and C are reported in Table 1.¹⁸ Due to potential endogeneity in specification B1, we favor specifications B2 and B3. When computing significance levels, we always use heteroscedasticity consistent covariances since heteroscedasticity tests (Sroeter, 1978; White, 1980; Cook and Weisberg, 1983) indicate heteroscedasticity in all cases. Testing for autocorrelation in panel data (Wooldridge, 2002; Drukker, 2003) on the other hand yields no clear indication for serial correlation in the residuals. However, since in some cases autocorrelation problems become obvious when regressing residuals on preceding residuals, we also use heteroscedasticity and autocorrelation consistent standard errors (Newey-West).

The F -tests for all coefficients jointly being zero reject the null hypothesis in all cases. However, the reported (adjusted) R^2 values are relatively low. One reason for the low explanatory power is that we estimate in first differences. The examination of the residuals' distributions reveals in all cases very high Jarque-Bera statistics mainly stemming from high Kurtosis values. Since this finding rejects the normal distribution

¹⁶ Note that including total income as well as per capita income simultaneously would lead to a multicollinearity problem.

¹⁷ Analog to specification B1, we also employ GDP lagged for one period as an instrument for current GDP in specification C1 following the result of the Durbin-Wu-Hausman test for endogeneity. Again, there is no indication for the presence of a weak instrument and the Sargan test does not suggest including GDP lagged for two periods as another instrument.

¹⁸ The results of the TSLS estimations of models B1 and C1 with GDP (per capita) instrumented by the corresponding lagged variable are not reported since all coefficients are insignificant. Endogeneity, especially of GDP (per capita), obviously is a caveat when interpreting the regression results. The results not reported here as well as standard errors and statistics of the various tests are available upon request.

assumption of the residuals, the reported significance levels should be interpreted with some caution. We could not remedy this problem by redefining the estimation model, changing the sample size or eliminating outliers.

Specification Method	A1 Country-RE	A2 FE	B1 FE	B2 FE lagged regressors	B3 FE lagged regressors and MA	C1 FE	C2 FE lagged regressors	C3 FE lagged regressors and MA
Observations	1630	1630	1563	1507	1439	1563	1509	1441
Countries	60	60	60	60	60	60	60	60
Years	1976-2004	1976-2004	1976-2004	1977-2004	1978-2004	1976-2004	1977-2004	1978-2004
Depend. var.	$\Delta \ln(EI)$	$\Delta \ln(EI)$	$\Delta \ln(EI)$	$\Delta \ln(EI)$	$\Delta \ln(EI)$	$\Delta \ln(EI)$	$\Delta \ln(EI)$	$\Delta \ln(EI)$
<i>CONST</i>	0.004	0.004 *	-0.012	-0.025 **	-0.024 **	0.001	0.016 *	0.014
ΔIND			0.005	0.027	-0.008	-0.001	0.102 *	0.098 *
<i>I</i>			0.146 ***	0.127 **	0.111 **	0.153 ***	0.068 **	0.062 *
<i>FDI / I</i>	-0.031 **	-0.027						
<i>FDI</i>			0.064	0.050	-0.042	0.072	0.196 ***	0.139
<i>IM</i>			-0.017	0.014	0.022	-0.010	0.002	0.007
<i>AID</i>			-0.030	-0.117 ***	-0.107 *	-0.024	0.010	0.029
$\Delta \ln(YPC)$			-0.788 ***	0.002	-0.004			
$\Delta \ln(Y)$						0.210 ***	0.078 **	0.086 **
Adj. R^2	0.003	0.071	0.408	0.076	0.070	0.191	0.126	0.105
<i>F</i> -stat.	5.973	2.414	12.571	2.352	2.190	4.972	3.358	2.861
Prob(<i>F</i> -stat.)	0.015	0	0	0	0	0	0	0

* Significant at the 10 % level, ** significant at the 5 % level, *** significant at the 1 % level; heteroscedasticity and autocorrelation consistent standard errors; Δ = first time differences; Country-RE = country-specific random effects; FE = country- and time-specific fixed effects; B3 and C3 use lagged regressors and FDI, IM, AID are moving averages of the past 3 years.

Table 1: Estimation results for specifications A, B and C

The coefficient of the variable of main interest, FDI inflows, is significant and negative in specification A1, which confirms the finding by Mielnik and Goldemberg (2002). In the slightly modified specification A2 with time- and country-specific fixed effects, the probability that the coefficient is different from zero already falls below any common significance level while the coefficient stays negative. In specification C2, where changes in total energy use are examined, the coefficient of lagged FDI inflows is positive and highly significant. According to the estimate, a one percentage point increase in FDI intensity raises total energy supply by about 0.2 percent. This finding is likely to stem from the scale effect: FDI inflows in the

previous year induce increasing economic activity in the current year, which results in higher energy use. All other regressions do not confirm any significant effect of FDI on energy intensity. Using our theoretical background, the results imply that it is not possible to identify a robust energy reducing effect of technology transfer via FDI in this macro panel.

We now compare the results for *FDI* with those for *IM* and *AID* and *I*. While imports do not show a significant effect in any regression, *AID* is significant and negative in specifications B2 and B3. This energy intensity reduction by aid inflows is in line with the expectations that industrialized donor countries promote energy saving technologies in developing countries. According to the results of B2, aid inflows amounting to one percent of GDP of the recipient country reduce the country's energy intensity by about 0.1 percent. On the contrary, the investment share in GDP labeled *I* is significantly positive in all specifications. Gross investment of one percent of GDP increases energy intensity or total primary energy supply by about 0.1 percent. The hypothesis that new capital investment brings about energy saving technical progress is therefore challenged. It is on the other hand possible that the investment variable absorbs part of the composition effect: Since energy intensive sectors are typically also capital intensive, a strong increase in investment may reflect an expansion of the energy intensive sectors and could therefore lead to an increase in energy intensity. This view is supported by the fact that our existing measure of the composition effect, namely the share of industrial value added in GDP, is rather crude and cannot capture all sectoral changes in the economy.

Regarding the remaining control variables, income per capita growth $\Delta \ln(YPC)$ is highly significant and negative in specification B1, but not in specifications B2 and B3 where it appears as a lagged variable. According to B1, a one percent increase in per capita income would reduce energy intensity by 0.79 percent. This strong effect is likely to stem from short-term GDP-fluctuations, where energy intensity, defined as energy supply over GDP, typically moves to the opposite direction than GDP. If a longer term influence really existed, we would find a significant result in specifications B2 or B3, but this is not the case. Furthermore, the significance of GDP per capita disappears also in specification B1 when the variable is instrumented by lagged GDP per capita as a robustness check. In models C1, C2 and C3, the change in total income $\Delta \ln(Y)$ has a significantly positive influence on total energy supply as expected (scale effect). The coefficient varies between 0.21 in

C1 and 0.08 in C2. Concerning the share of industrial value added in GDP, we always find a positive sign, but the coefficient is significant only in specifications C2 and C3. It is likely that the sectoral change between the industry sector on the one hand and agriculture or services on the other hand is less important than sectoral changes within the industry sector. Unfortunately, no such detailed data are available for the countries in our sample, so it is also likely that part of the composition effect is implicitly included in the coefficient of FDI.

We also extend the basic model (specification B) in order to investigate the interaction of FDI with country-specific characteristics. We add interaction terms of the aggregate FDI inflow intensity with changes in the share of industrial value added in GDP, with the import intensity and with shares of energy sources in total primary energy supply (particularly coal, oil, gas, nuclear power and hydro power). The interaction of the FDI inflow intensity with the industry share in GDP examines whether FDI inflows coming along with changes in the sectoral composition influence the energy intensity in a country differently. Note that this interaction term includes the composition effect of FDI in the case of a systematic influence of FDI inflows on the industry share. The interaction of the FDI inflow intensity with the import intensity examines whether FDI and imports jointly affect the energy intensity due to increasing intra-firm imports by multinational enterprises. The interaction of the FDI inflow intensity with shares of energy sources determines whether the potential of FDI to affect energy intensity is related to the energy mix of a country. We run several regressions, including only one interaction term at the same time.¹⁹ We find no evidence for an energy reducing effect of the aggregate FDI inflow intensity, and none of the interaction terms leads to significant results. Nevertheless, the interaction term analyses show a way for further research, focusing more on the interactions of FDI with other economic indicators.

¹⁹ When adding the interaction term “FDI multiplied with the share of the energy source” we also add the share of the energy source as a separate regressor.

4 Robustness checks

In the robustness checks we address the following concerns: Technological improvements can foremost be expected from greenfield investment, i.e. installations of new production facilities and machinery, and not so much from pure ownership changes. Hence we correct our FDI data by subtracting mergers and acquisitions in part 4.1. It would also be interesting to distinguish market seeking horizontal FDI from low production cost seeking vertical FDI, but such an empirical analysis is not possible due to a lack of data. Another reason for the inability to find any robust influence of FDI on energy intensity might be the fact that we use aggregated data without a distinction between sectors and between source countries of FDI. While FDI flows into energy intensive sectors might have a significant potential for reducing economy-wide energy use, this might not be the case for FDI flows into other sectors. Furthermore, FDI from the U.S. might embody different technologies than FDI from Germany or Japan and therefore lead to other effects on energy use. In order to address these issues, we introduce new sectoral data for the U.S. as a source country of FDI in developing countries in part 4.2. Unfortunately, the robustness checks cannot be scrutinized with the full original data sample because of insufficient data availability. The price we have to pay is a reduced or modified sample in both robustness checks.

4.1 M&A correction of FDI

To get an approximation for greenfield investment, we subtract values of mergers and acquisition (M&A) from the FDI values we used before.²⁰ The M&A data are taken from UNCTAD (2008), the other data are the same as before. Due to a lack of data in the M&A database, the original sample is now reduced to the time frame of 1987 to 2004 and to 58 countries.²¹

In general, the results confirm those found before (see Table A5 in the Online Appendix at <http://journals.cambridge.org/EDE>). In none of the specifications, there

²⁰ This is no exact way of computing greenfield FDI values, but probably the best available approximation.

²¹ Without Benin and Togo.

is a significant energy reducing effect of FDI. As with the original dataset, we find a significantly positive effect of the FDI intensity on the growth rate of total energy use in specification C2, which arises from the scale effect of FDI. The reduction of the growth rate of energy intensity via aid inflows can be confirmed in specifications B1, B2 and B3. Aid inflows of one percent of GDP reduce energy intensity by about 0.06 to 0.19 percent. The coefficient of aid is also significantly negative in C1 and C3. In contrast to the former regressions with the larger dataset, imports turn out to increase energy intensity in specifications B2 and B3. However, this effect is only weakly significant, both statistically and economically.

4.2 Sectoral data for the United States as a source country

Finally, we look specifically at FDI outflows from the United States and apply data from the Bureau of Economic Analysis (2008). Due to restricted data availability for the recipient countries we need to narrow the sample to 10 developing countries²² and four sectors²³ from 1982 to 2003. Sectoral value added data are taken from UNIDO (2006) and energy data are IEA data as before. Both energy intensity and FDI in each country's sector are measured relative to the value added in the respective sector. In order to obtain consistent estimates for energy intensity, the value added data are converted into constant PPP using the Penn World Table (Heston et al., 2006) and a U.S. consumer price index (International Monetary Fund, 2007b). We employ a parsimonious model specification in first time differences with the logarithm of energy intensity as the dependent variable and FDI as the only explanatory variable. In the second specification, we employ one-year lagged FDI values as the explanatory variable. We estimate a system of four equations, one for each sector, and apply the technique of Seemingly Unrelated Regressions (SUR). This technique has the advantage that we gain separate results for each of the sectors. At the same time, it exploits the common information in the residuals across sectors and is therefore potentially more efficient than four separate regressions. We include a constant and

²² The countries are: Argentina, Chile, China, Colombia, India, Indonesia, Mexico, Philippines, South Africa, Venezuela.

²³ The four sectors are on a 3-digit ISIC level: chemical products (industrial and other chemicals without petrochemical industry); food, beverages and tobacco; machinery (non-electrical and electrical); metals (including iron and steel and non-ferrous metals).

estimate a pooled regression without fixed effects since the *F*-/*LR*-tests do not reject the null hypothesis of redundant fixed effects.

In the results (see Table A6 in the Online Appendix at <http://journals.cambridge.org/EDE>), the coefficient of FDI inflows is not statistically significant in any of the sectors and once again, the hypothesis of energy saving technology transfer via FDI is not supported. Note that one reason for the inability to identify an energy reducing technology effect might be a composition effect of FDI towards more energy intensive products within the still aggregated sectors chemicals, food, machinery and metals.

5 Conclusion

Referring to the hypothesis of energy saving technology transfer proposed by Mielnik and Goldemberg (2002), this paper examines whether foreign direct investment inflows reduce energy intensities of developing countries. Such reductions are desirable against the background of climate change mitigation since energy use is strongly related to emissions of carbon dioxide. Theoretically, the impact of FDI inflows can be decomposed into a scale, a composition and a technique effect (Grossman and Krueger, 1993). Using the theoretical principles provided by Antweiler et al. (2001) and applying the concept to energy use, we set up an empirical macro-level model in order to identify the effects of technology transfer via FDI on energy intensities in developing countries.

We investigate cross-country panel data on 60 developing countries for the years 1975 to 2004. In the results, we find no energy reducing effect of FDI inflows. Thus we cannot confirm the hypothesis of energy saving technology transfer via FDI by Mielnik and Goldemberg (2002) in general. Regarding the influence of foreign aid inflows on energy intensity, we find a significantly negative effect in several specifications. In order to take country-specific characteristics into account, we also examine the interaction of FDI inflows with changes in the industry share, with imports and with shares of energy sources (particularly coal, oil, gas, nuclear power and hydro power). We find no evidence for a significant joint effect of FDI inflows together with the other determinants on energy intensity. Finally, we carry out robustness checks with modified and new data sets. Correcting FDI inflow values by

data on mergers and acquisitions as an approximation for greenfield investment confirms in general the original results. While energy savings via FDI cannot be found, the energy efficiency gains from foreign aid are confirmed. Focusing on the U.S. as a single source country and using sectorally disaggregated data, we find again no indication for energy savings via FDI.

However, there are caveats when interpreting the regression results. At first, the explanatory power of our model is low. On the one hand, this is caused by the loss of information due to differencing the estimating equation; on the other hand, the energy intensity of a country is determined by many technical, infrastructural, economic and political factors that cannot all be captured in a macro model, because the necessary data are not available. To consider as many unobservable effects as possible, we use panel data models with time- and country-specific effects. The second caveat is the high kurtosis of the residual distribution that biases the standard errors and consequently the significance levels. Reducing outliers from the sample and different model variants could not remedy this problem. If we restricted the study on time series of single countries, the vast heterogeneity in the sample would be reduced, and the properties of the residual distribution would be more favorable. Then, however, it would no longer be possible to derive general results. Finally, it is possible that while some FDI might reduce energy intensity via technology transfer, other FDI might induce a shift towards more energy intensive production via a change in the sectoral composition of production. The significance of the results thus possibly suffers from the fact that we cannot fully disentangle technology transfer from the composition effect.

Nevertheless, our results suggest that energy efficiency gains in developing countries via FDI do not occur automatically and without climate or energy policy. Thus, it is sensible to make use of issue-linkage, as for instance intended by the clean development mechanism of the Kyoto Protocol (for technology transfer in CDM projects see Seres, 2007) or international energy technology collaboration (Justus and Philibert, 2004) to explicitly encourage foreign direct investment that brings along energy reducing technology transfer. A technology fund governed by the World Bank may also help to direct investment into energy saving projects in developing countries. Note furthermore that our results could change in the presence of a carbon price in the countries under scrutiny since pricing carbon creates an incentive to invest into energy saving technologies.

A challenge for further research is to identify country-specific characteristics that enhance technology transfer via FDI. This can be accomplished by using interaction terms of FDI and country-specific variables. Departing from the panel analysis, it can be helpful to use data on specific countries in order to analyze whether FDI inflows reduce energy intensity in these specific cases. There is a broad literature on technology spillovers in general, but technology spillovers affecting energy intensity have not been investigated in depth. One can include long-run effects to examine whether spillovers from FDI decrease energy intensity with a time delay. Where sectoral data is available, the analysis can be performed in a more detailed way and one can try to distinguish explicitly the composition effect from the technology transfer. A better understanding of international technology diffusion in the context of climate change is highly important for supporting developing countries in achieving lower energy and carbon intensities than the industrialized countries did in the past.

References

- Aitken, B. J. and A. E. Harrison (1999). Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela, *The American Economic Review* 89(3), 605-618.
- Antweiler, W., B. R. Copeland and M. S. Taylor (2001). Is Free Trade Good for the Environment?, *The American Economic Review* 91(4), 877-908.
- Arrow, K. J. (1962), The Economic Implications of Learning by Doing, *The Review of Economic Studies* 29(3), 155-173.
- Bureau of Economic Analysis (2008). U.S. direct investment abroad, Online database. <http://www.bea.gov/international/>.
- Birol, F. and B. E. Okogu (1997). Purchasing-Power-Parity (PPP) Approach to Energy-Efficiency Measurement: Implications for Energy and Environmental Policy, *Energy* 22 (1), 7-16.
- Choi, I. (2001). Unit Root Tests for Panel Data, *Journal of International Money and Finance* 20, 249-272.
- Cole, M. A. (2006). Does trade liberalization increase national energy use?, *Economics Letters* 92, 108-112.
- Copeland, B. R. and M. S. Taylor (2003). *Trade and the Environment: Theory and Evidence*, Princeton University Press, Princeton.

- Cook, R. D. and S. Weisberg (1983). Diagnostics for heteroscedasticity in regression, *Biometrika* 70, 1-10.
- Corcos, G., M. Del Gatto, G. Mion and G. I. P. Ottaviano (2007). Productivity and Firm Selection: Intra- vs. International Trade, CORE Discussion Paper No. 2007/60, Université catholique de Louvain.
- Dickey, D. A. and W. A. Fuller (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association* 74, 427-431.
- Drukker, D. M. (2003). Testing for serial correlation in linear panel-data models, *Stata Journal* 3(2), 168-177.
- Eskeland, G. and A. Harrison (2003). Moving to greener pastures? Multinationals and the pollution haven hypothesis, *Journal of Development Economics* 70(1), 1-23.
- Fisher-Vanden, K., G. H. Jefferson, H. Liu and Q. Tao (2004). What is driving China's decline in energy intensity?, *Resource and Energy Economics* 26, 77-97.
- Galli, R. (1998). The Relationship Between Energy Intensity and Income Levels: Forecasting Long Term Energy Demand in Asian Emerging Countries, *The Energy Journal* 19(4), 85-105.
- Granger, C. W. J. and P. Newbold (1974). Spurious regressions in econometrics, *Journal of Econometrics* 2, 111-120.
- Grossman, G. and A. Krueger (1993), 'Environmental Impacts of a North American Free Trade Agreement', in P. Garber (ed.), *The U.S.-Mexico free trade agreement*, MIT Press, Cambridge, MA, pp. 13-56.
- Heston, A., R. Summers and B. Aten (2006), *Penn World Table Version 6.2*, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006.
- Im, K. S., M. H. Pesaran and Y. Shin (2003). Testing for unit roots in heterogeneous panels, *Journal of Econometrics* 115, 53-74.
- International Energy Agency (2007). *World Energy Statistics and Balances*, Online database: <http://data.iea.org/ieastore/statslisting.asp>.
- International Monetary Fund (2007a). *Balance of Payments Statistics*, Online database: <http://www.imfstatistics.org/bop/>.
- International Monetary Fund (2007b). *International Financial Statistics*, Online database: <http://www.imfstatistics.org/imf/>.
- Javorcik, B. S. and M. Spatareanu (2008). To share or not to share: Does local participation matter for spillovers from foreign direct investment?, *Journal of Development Economics* 85(1-2), 194-217.
- Javorcik, B. S. (2004). Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages, *The American Economic Review* 94(3), 605-627.

- Justus, D. and C. Philibert (2005). International Energy Technology Collaboration and Climate Change Mitigation, Synthesis Report, OECD/IEA, Paris.
- Keller, W. (2004). International Technology Diffusion, *Journal of Economic Literature* 42, 752-782.
- Keller, W. and S. R. Yeaple (2009). Multinational Enterprises, International Trade, and Productivity Growth: Firm-Level Evidence from the United States, Forthcoming in *The Review of Economics and Statistics*.
- Levin, A., C.-F. Lin and C.-S. J. Chu (2001). Unit root tests in panel data: asymptotic and finite sample properties, *Journal of Econometrics* 108, 1-24.
- Maddala, G. S. and S. Wu (1999). A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test, *Oxford Bulletin of Economics and Statistics* 61(1), 631-652.
- Mazzanti, M., A. Musolesi and R. Zoboli (2006). A Bayesian Approach to the Estimation of Environmental Kuznets Curves for CO₂ Emissions, *Nota di lavoro* 121.06, Fondazione Eni Enrico Mattei, Milan.
- Mielnik, O. and J. Goldemberg (2000). Converging to a common pattern of energy use in developing and industrialized countries, *Energy Policy* 28, 503-508.
- Mielnik, O. and J. Goldemberg (2002). Foreign direct investment and decoupling between energy and gross domestic product in developing countries, *Energy Policy* 30, 87-89.
- Perman, R. and D. I. Stern (2003). Evidence from panel unit root and cointegration tests that the Environmental Kuznets Curve does not exist, *The Australian Journal of Agricultural and Resource Economics* 47(3), 325-347.
- Peterson, S. (2008). Greenhouse gas mitigation in developing countries through technology transfer?: a survey of empirical evidence, *Mitigation and Adaptation Strategies for Global Change* 13(3), 283-305.
- Phillips, P. C. B. and P. Perron (1988). Testing for Unit Root in Time Series Regression, *Biometrika* 75, 335-345.
- Saggi, K. (2002). Trade, Foreign Direct Investment, and International Technology Transfer: A Survey, *The World Bank Research Observer* 17(2), 191-235.
- Seres, S. (2007). Analysis of Technology Transfer in CDM Projects, Report prepared for the UNFCCC Registration and Issuance Unit CDM/ SDM.
- Siebert, H. (2005). *Economics of the Environment – Theory and Policy*, Sixth Edition, Springer-Verlag, Heidelberg.
- Szroeter, J. (1978). A Class of Parametric Tests for Heteroscedasticity in Linear Econometric Models, *Econometrica* 46, 1311-28.
- Stern, D. I. (2004). The Rise and Fall of the Environmental Kuznets Curve, *World Development* 32(8), 1419-1439.

Tybout J. R. (2002). Plant and Firm-Level Evidence on New Trade Theories, in J. Harrigan (ed.), Handbook of International Economics 38, Basil-Blackwell.

UNCTAD (2008). FDIstat, Online database: <http://stats.unctad.org/FDI/>.

UNIDO (2006). INDSTAT3, Industrial Statistics Database at the 3-digit level of ISIC (Revision 2), CD-ROM.

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, *Econometrica* 48, 817-838.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, MA.

World Bank (2007). World Development Indicators, Online database: <http://go.worldbank.org/B53SONGPA0>.

Appendix

Variable	Definition	Unit
<i>E</i>	Total primary energy supply	ktoe
<i>EI</i>	Energy intensity in purchasing power parity (total primary energy supply divided by GDP in PPP)	ktoe per (constant 2000) million dollar in PPP
<i>IND</i>	Share of industrial value added in GDP	-
<i>I</i>	Gross fixed capital formation as a share of GDP	-
<i>FDI / I</i>	Net inflows of foreign direct investment as a share of gross fixed capital formation	-
<i>FDI</i>	Net inflows of foreign direct investment as a share of GDP	-
<i>IM</i>	Imports as a share of GDP	-
<i>AID</i>	Official Development Assistance and Official Aid inflows as a share of GDP	-
<i>Y</i>	Total income (measured by GDP in PPP)	(constant 2000) million dollars in PPP
<i>YPC</i>	Per capita income (measured by GDP in PPP)	(constant 2000) dollars in PPP per capita

Table 2: Definition of variables

Section 2.2

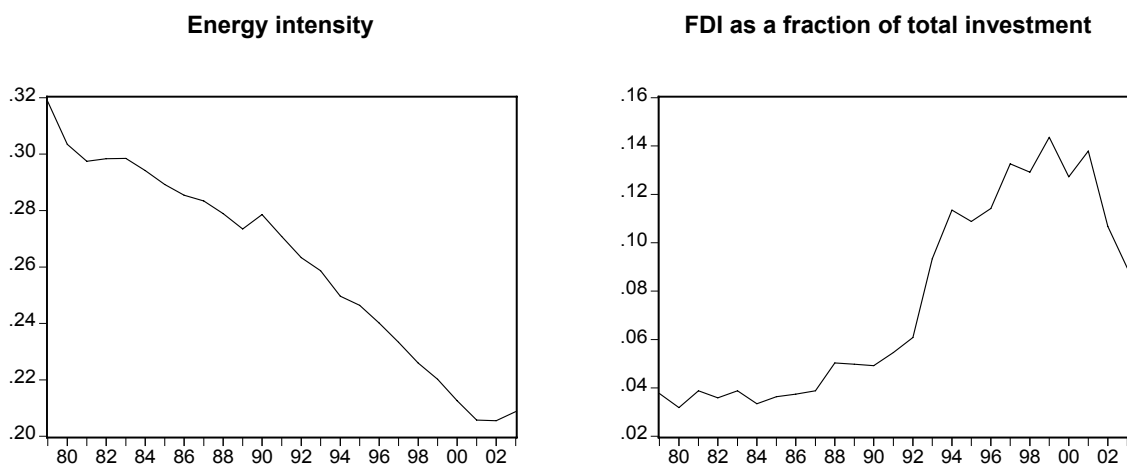


Figure A1: Trends of the variables in the simplified model with 20 countries, 1979-2003

Energy Intensity			
In levels		in 1st differences	
With constant and linear trend	-1.76 (-3.61)	with constant	-5.39 (-3.00)
Only with constant	-0.64 (-2.99)	no constant	-5.43 (-1.96)
FDI as a fraction of total investment			
In levels		in 1st differences	
With constant and linear trend	-0.96 (-3.61)	with constant	-9.73 (-3.00)
Only with constant	-1.01 (-2.99)	no constant	-9.85 (-1.96)

In parentheses: 5% critical values to reject the unit root null hypothesis.

Table A1: Unit root test statistics (Augmented Dickey Fuller) for the variables in the simplified model

Model Specification	Trace statistic
Intercept in CE, none in VAR	16.37 (20.26)
Intercept in CE and VAR	7.10 (15.49)
Intercept and trend in CE, intercept in VAR	20.90 (25.87)

In parentheses: 5% critical values to reject the null of no cointegration.

Table A2: Johansen cointegration test of the variables in the simplified model

Section 3: Detailed data description

Table A3 gives an overview on the data that is used in our empirical estimations. The large differences between the minimal and maximal values of the variables in the sample indicate an obvious heterogeneity of countries and years. For instance, per capita income *YPC* ranges from about 485 to 23,266 dollars in PPP. The heterogeneity of countries is also apparent from the rising, falling and undefined time trends of energy intensity and foreign direct investment. Figures A2 and A3 visualize that in graphs of four typical countries of the sample. It is also noteworthy that some countries exhibit negative FDI inflows in certain years. This can for instance be the case when foreign companies withdraw from the market or disinvest.

Variable	<i>E</i>	<i>EI</i>	<i>IND</i>	<i>I</i>	<i>FDI/I</i>	<i>FDI</i>	<i>IM</i>	<i>AID</i>	<i>Y</i>	<i>YPC</i>
Obs.	1794	1750	1696	1700	1688	1745	1741	1733	1756	1756
Mean	42,565	0.257	0.331	0.215	0.082	0.017	0.346	0.041	176,962	4,581
Min.	276	0.040	0.062	0.021	-1.389	-0.122	0.030	-0.007	955	485
Max.	1,609,348	1.176	0.775	0.606	1.983	0.401	1.075	0.956	7,023,283	23,266
Std. dev.	128,874	0.182	0.116	0.069	0.147	0.028	0.184	0.069	486,419	3,622

Obs. = number of available observations, which differs between variables; std. dev. = standard deviation.

Table A3: Descriptive statistics of the sample of 60 developing countries in 1975-2004

Energy Intensity

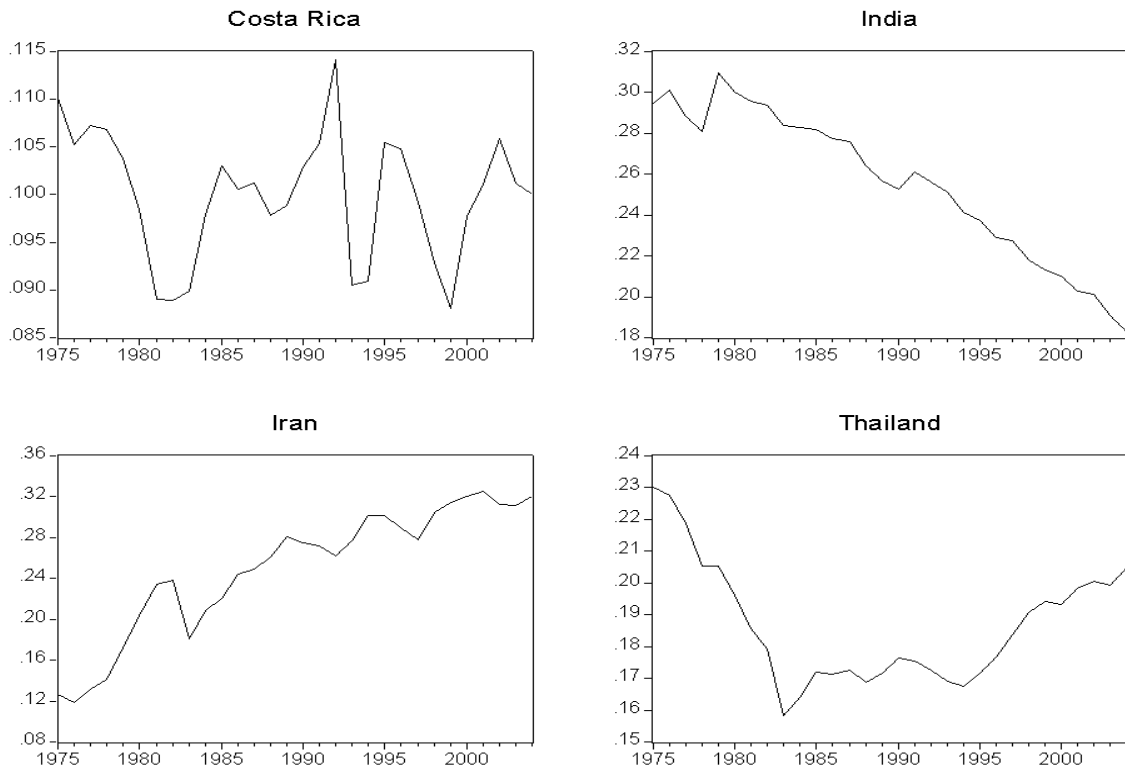


Figure A2: Time trends of energy intensity 1975-2004 in four countries of the sample

FDI Inflow

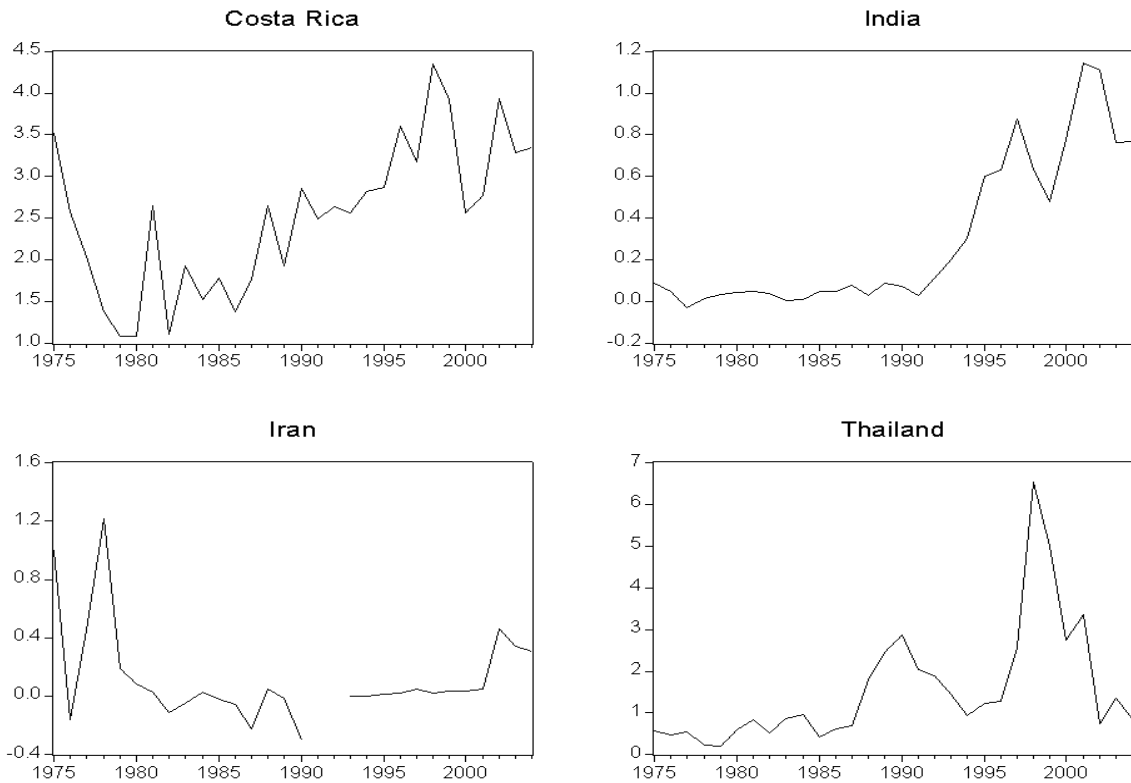


Figure A3: Time trends of FDI in percent of GDP 1975-2004 in four countries of the sample

The following paragraphs describe obvious trends of important variables revealed by a closer look at the distinct time series.

As expected, GDP (in PPP), denoted by Y , rose during this period in all countries. A number of countries show a continuous increase, for example China, India and Pakistan. While India's and Pakistan's GDP expanded during these 30 years by a factor of almost 5, China's GDP exploded by a factor of 13. However, there are economies with tremendous GDP fluctuations such as Nicaragua or Peru. In many countries YPC , income per capita, grew in a similar way as total GDP, but this is not necessarily the case. Nicaragua and Venezuela for instance show a falling trend, other countries' income per capita fluctuated around a constant level.

Since increasing production and consumption reflected by GDP growth lead to higher energy demand, it is not surprising that total primary energy supply E clearly rose in all countries as well. While most countries' total energy use increased in a smooth continuous way, some countries like Peru and Uruguay show large fluctuations. Such energy use and GDP changes might stem from political disturbances or other economic shocks, which cannot be captured in the econometric analysis and will probably create estimation errors. The development of energy intensities EI is ambiguous across countries. In 17 countries (including China, India and Peru) energy intensity declined. China's energy intensity, starting from a high level in 1975, decreased by approximately 75 % until 2004. On the other hand, energy intensity increased in 21 countries (Algeria, Saudi Arabia, Iran and others). 22 countries show no obvious tendency. Having a closer look at the time series of energy use E and GDP Y , it becomes obvious in a number of cases that GDP short-time fluctuations or shocks do not correspond with proportional fluctuations of energy supply. In these cases, GDP jumps up or down while energy supply is sluggish. As a consequence energy intensity defined as E / Y moves to the opposite direction of the GDP fluctuation. This is an important observation, which has to be taken into account in the specification of the estimation models. (Another question not discussed here is the role of measurement errors and measurement difficulties.)

A number of countries (such as Bangladesh, Honduras, India, Indonesia, Republic of Korea, Senegal and Thailand) show a clearly rising share of the industry sector IND over time. Other countries (like Argentina, Oman and Zambia) have falling shares. In the remaining cases IND fluctuates or shows upward or downward trends within the time frame of the sample. In China the industry share remained relatively stable over time, reaching its

maximum of 48.2 % relative to GDP in 1978 and its minimum of 41.6 % in 1990, while the 2004 share was 46.2 % of GDP.

Gross investment in absolute terms had an upward sloping tendency, and several economies had a stable continuing increase in investment during the sample period (Chile, China, India, Korea, Pakistan and others). This upward trend vanishes in many cases when looking at I , gross investment relative to GDP. Investments in China rose from 39.2 billion USD (29.4 % of GDP) in 1975 to 658.2 billion USD (38.4 % of GDP) in 2004.

Net FDI inflows (In US Dollars referring to the year 2000) show a rising tendency in most countries, especially during the 1990s. FDI relative to GDP, here labeled FDI , also rose in many countries, but this trend is less obvious than the increase in absolute FDI inflows. Some countries show periods with high fluctuations or plummeting FDI shares. FDI inflows to China increased strongly from 51 million USD (0.03 % of GDP) in 1980 to 48.7 billion USD (2.8 % of GDP) in 2003. The highest ratio of FDI to GDP (in other words the highest intensity of FDI inflows) was reached in 1993 with 6.3 %. When dividing FDI inflows by gross investment instead of GDP, the resulting time series data are very similar.

Besides FDI, imports are another indicator for the integration of a country into the world economy and a potential channel for technology transfer. Their value had an upward sloping trend between 1975 and 2004 in all countries, while periods of decline or years of plummeting imports occurred in some countries. The imports relative to a country's GDP, denoted by IM , clearly rose only in 23 countries. The other countries show decreases in import intensities or fluctuations. China's import value increased from roughly 6.1 billion USD (4.6 % of GDP) in 1975 to 538.5 billion USD (31.4 % of GDP) in 2004. The reception of international aid is a further potential source of international technology transfer. In contrast to FDI and trade, there is no clear trend of aid inflows when examining the time series of the 60 countries in the sample. Referring to aid inflows relative to GDP, called AID in the data set, some countries show an upward trend (e.g. Ghana) and other countries have falling aid intensities (e.g. India, Tunisia). Absolute and relative aid flows to China reached their maximum in 1993 and declined in the following years.

We conclude that there are increasing time trends of energy supply, imports and foreign as well as gross investment in accordance with GDP growth. It is difficult to observe any direct relationship between energy and these variables besides the time trend. When looking at intensities, i.e. the variables divided by GDP, a different picture with considerable heterogeneity of the 60 countries arises. An econometric analysis of this panel data may reveal whether FDI has a significant influence on energy intensity. Figure A4 visualizes the

facts about the Chinese economy discussed before. Obviously FDI and international trade have played an increasingly important role. However, this effect is weaker or not detectable in other developing countries.

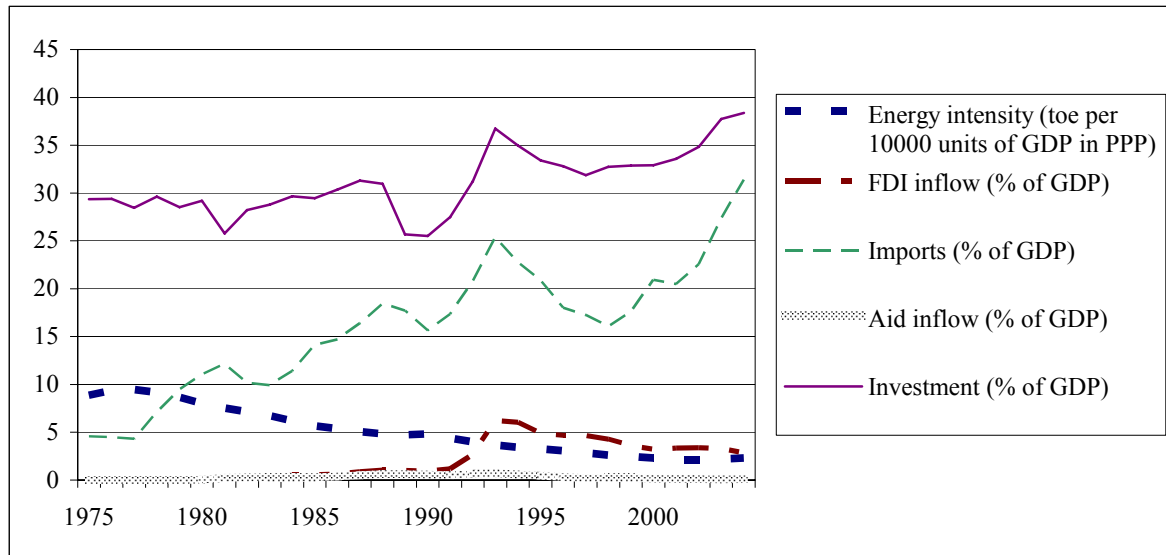


Figure A4: Indicators of the Chinese economy, 1975-2004

Section 3.2

	ΔIND	I	FDI	IM	AID	$\Delta \ln(YPC)$	$\Delta \ln(Y)$
ΔIND	1						
I	0.016	1					
FDI	0.063	0.135	1				
IM	0.050	0.351	0.350	1			
AID	0.031	-0.069	0.022	0.184	1		
$\Delta \ln(YPC)$	0.166	0.275	0.113	0.149	-0.018	1	
$\Delta \ln(Y)$	0.167	0.276	0.088	0.156	0.020	0.984	1

Table A4: Pair wise correlations of the explanatory variables

Section 4.1

Specification Method	A1 Country-RE	A2 FE	B1 FE	B2 FE lagged regressors	B3 FE lagged regressors and MA	C1 FE	C2 FE lagged regressors	C3 FE lagged regressors and MA
Observations	956	956	924	869	804	924	871	1441
Countries	58	58	58	58	58	58	58	60
Years	1988-2004	1988-2004	1988-2004	1989-2004	1990-2004	1988-2004	1989-2004	1978-2004
Depend. var.	$\Delta \ln(EI)$	$\Delta \ln(EI)$	$\Delta \ln(EI)$	$\Delta \ln(EI)$	$\Delta \ln(EI)$	$\Delta \ln(EI)$	$\Delta \ln(EI)$	$\Delta \ln(EI)$
<i>CONST</i>	0.000	0.000	-0.001	-0.017	-0.031	0.011	0.038 ***	0.023 *
ΔIND			-0.029	0.302 ***	0.208 *	-0.030	0.172 ***	0.122 **
<i>I</i>			0.099 **	-0.038	0.033	0.102 **	-0.013	0.027
<i>FDI / I</i>	-0.022	-0.024						
<i>FDI</i>			0.045	-0.025	-0.217	0.044	0.143 *	0.032
<i>IM</i>			-0.019	0.081 *	0.091 *	-0.012	-0.010	0.018
<i>AID</i>			-0.061 **	-0.117 ***	-0.185 **	-0.054 **	-0.039	-0.107 **
$\Delta \ln(YPC)$			-0.746 ***	0.120	0.182 **			
$\Delta \ln(Y)$						0.251 ***	0.057 *	0.084 **
Adj. R^2	0.003	0.042	0.380	0.091	0.090	0.139	0.075	0.067
<i>F</i> -stat.	3.287	1.561	8.168	2.352	2.035	2.881	1.899	1.752
Prob(<i>F</i> -stat.)	0.070	0.002	0	0	0	0	0	0

* Significant at the 10 % level, ** significant at the 5 % level, *** significant at the 1 % level; heteroscedasticity and autocorrelation consistent standard errors; Δ = first time differences; Country-RE = country-specific random effects; FE = country- and time-specific fixed effects; B3 and C3 use lagged regressors while FDI, IM and AID are moving averages of the past 3 years.

Table A5: Estimation results for specifications A, B and C with M&A corrected FDI

Section 4.2

Method	Seemingly Unrelated Regression pooled				Seemingly Unrelated Regression pooled, lagged FDI values as regressor			
System Obs.	454				456			
Countries	10				10			
Years	1983-2003				1983-2003			
Depend. var.	$\Delta \ln(EI)$				$\Delta \ln(EI)$			
Sector	Chem.	Food	Mach.	Metals	Chem.	Food	Mach.	Metals
Observations	149	130	44	131	150	128	48	130
<i>CONST</i>	-0.017	-0.026 *	-0.014	-0.004	-0.020	-0.017	-0.011	-0.007
<i>FDI</i>	-0.064	0.276	0.328	-3.235	0.240	-0.645	0.491	3.104
R^2	0.001	0.001	0.000	0.006	0.001	0.005	0.004	0.017

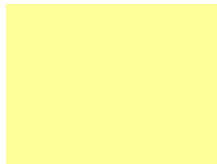
* Significant at the 10 % level, ** significant at the 5 % level, *** significant at the 1 % level;

Δ = first time differences, Chem. = Chemicals, Mach. = Machinery

Table A6: Estimation results for the sectoral dataset with the United States as source country

Chapter 4:
Can Carbon Based Import Tariffs Effectively Reduce Carbon Emissions?

Kiel Working Paper No. 1565
<http://www.ifw-kiel.de/pub/kap>



Chapter 4:
Can Carbon Based Import Tariffs Effectively Reduce
Carbon Emissions? *

Michael Hübler †

October 20, 2009

Abstract

We estimate CO₂ implicitly contained in traded commodities based on the GTAP 7 data: While net carbon imports into the industrialized countries amount to 15% of their total emissions, net carbon exports of the developing countries amount to 12% of their total emissions, and net carbon exports of China amount to 24% of China's total emissions. We also analyze policies under a global per capita emissions based contraction and convergence regime with emission trading: When China joins the regime, the developing countries will benefit, while the industrialized countries will be almost unaffected. When China does not join the regime and instead a carbon content based border tax is imposed, the industrialized countries will significantly benefit, while China will be significantly worse off. The effect of the border tax adjustment on the global carbon price and on global emissions seems negligible.

JEL Classifications: F13, F18, Q54

Keywords: Carbon content of trade, border tax adjustment, climate policy,
contraction and convergence, China

*I thank Sonja Peterson, Johannes Bröcker, Till Requate, Daiju Narita and Matthias Weitzel. We thank the Leibniz Association (WGL) for supporting our work.

†Kiel Institute for the World Economy, Düsternbrooker Weg 120, 24105 Kiel, Germany, Email: michael.huebler@ifw-kiel.de, Tel: +49-431-8814-401.

1 Introduction

The necessary drastic reduction in global CO₂ emissions critically depends on the inclusion of the developing and emerging economies, especially of China. - If China stays reluctant to join a binding post-Kyoto regime, China's emissions can possibly be reduced by imposing a border tax based on the carbon content of the traded commodities, since China is a major exporter of commodities. If such a border tax is imposed, China might react by imposing import tariffs as well. The question is how such policies affect global emissions or the global carbon price, and welfare given a certain climate policy scenario such as a contraction and convergence regime.

In this context, Paul Krugman writes in his New York Times blog under the title "The WTO is making sense" (Krugman 2009):

"There was some question about how the WTO would handle cap-and-trade whether it would accept the need for carbon tariffs, if some countries (cough China cough) drag their feet, or whether it would adopt a purist free-trade rule. The answer seems to be in - the WTO is going to treat cap-and-trade the same way it treats VATs, with border taxes allowed if they can be seen as reducing distortions.

One way to think about this is to say that the price of emissions licenses is ultimately a tax on consumers - and consumers should pay the same tax on emissions tied to imports as they do on emissions tied to domestic production. (That's the same reason you can charge VAT on imports.)

The same logic would also suggest that export subsidies are OK, but from an environmental point of view they're a bad idea; more broadly, the WTO view doesn't really take on the problem of negative externalities generated by foreigners producing for themselves.

But still, a sensible judgment."

The New York Times writes under the title "Possible Plan for Tariffs on Imports From China Remains Alive in House Climate Bill" on its web site (Friedman 2009):

"A House committee working on sweeping energy legislation seems determined to make sure that the United States will tax China and other carbon polluters, potentially disrupting an already-sensitive climate change debate in Congress. The Ways and Means Committee's proposed bill language would virtually require that the president impose an import tariff on any country that fails to clamp down on greenhouse gas emissions. ...

But associations that represent importers and multinational corporations are raising red flags, warning that the language could lead to trade wars, hurt the United States' ability to export low-carbon technology and harm consumers."

Moreover, The New York Times writes under the title "Obama Opposes Trade Sanctions in Climate Bill" (Broder 2009):

”’At a time when the economy worldwide is still deep in recession and we’ve seen a significant drop in global trade,’ Mr. Obama said, ‘I think we have to be very careful about sending any protectionist signals out there.’ He added, ‘I think there may be other ways of doing it than with a tariff approach.’”

From an economic point of view, the topic ”carbon content of trade” is related to the issue of carbon leakage. This issue has been extensively investigated in the literature - with diverse results and conclusions.¹ Following Marschinski et al. (2009), carbon leakage can occur through three channels: ”(1) free-rider leakage, i.e. lower incentives to contribute to the provision of a public good (environmental quality) as a strategic response to another actor’s effort; (2) specialisation leakage, i.e. relocation of production of energy-intensive goods due to changes in relative prices; (3) supply-side leakage, where decreased demand in one region leads to drops in fossil fuel prices and therefore results in higher consumption in other parts of the world.” Our CGE (computable general equilibrium) analysis will capture the second and the third channel.

Several studies have recently estimated carbon emissions implicitly embodied in traded commodities for different countries and specifically for China.² Shui and Harriss (2006) estimate that US CO₂ emissions would be 3 to 6% higher if the goods imported from China were produced in the USA, and that 7 to 14% of China’s CO₂ emissions can be attributed to exports for US consumers.

Peters and Hertwich (2008) calculate carbon contents of trade based on the GTAP 6 data set for 2001. (Herein, net carbon exports mean implicit CO₂ exports via exports of commodities minus implicit CO₂ imports via imports of commodities.) They find net carbon imports for the Annex B region of 5.6% relative to total CO₂ emissions produced in this region, and relative net carbon exports of 8.1% for the non-Annex B region. In particular, according to their calculations China’s net carbon exports amount to 17.8% of its total produced emissions, US net carbon imports amount to 7.3%, Japan’s to 15.3%, and Germany’s to 15.7%. Switzerland (122.9%) and Latvia (60.7%) are the most intensive net carbon importers among Annex B countries, while Hong Kong (182.2%), the rest of South African CU (176.4%) and Mozambique (172.4%) are the main net carbon importers among all countries. South Africa (38.2%) and the Russian Federation (21.6%) are the most intensive net carbon exporters among all countries.

Pan et al. (2008) estimate China’s emissions in 2006 on a consumption basis amount-

¹Compare for example IPCC (2007). Sijm et al. (2004) provide a detailed study. Marschinski et al. (2009) provide a recent review.

²For a ”review of input-output models for the assessment of environmental impacts embodied in trade” see Wiedmann et al. (2007). For an overview of quantitative analyses of CO₂ embodiment in international trade see Liu and Wang (2009).

ing to 3.8Gt of CO₂ rather than 5.5Gt on the standard production basis. This implies that China's net carbon exports amount to 1.7Gt in 2006.

These results emphasize the relevance of consumption based emissions accounting as a policy option that takes implicit international carbon trade into account. Also, in the absence of a global carbon price as the first best solution, carbon based border tax adjustments could internalize the negative external effects of carbon emissions as a second best solution.

But a number of authors, such as Bhagwati and Mavroidis (2007), question the economic, juristic and political feasibility of carbon content based border tax adjustment (BTA).

In the CGE model based literature on border tax adjustment, possible competitiveness disadvantages for firms within the European emissions trading scheme towards non-EU firms play a central role. Alexeeva-Talebi et al. (2008a) compare border tax adjustment based on imported *quantities* multiplied by domestic carbon content factors within an integrated emissions trading scheme based on imported *emissions* created during the production of imported commodities. They conclude that border tax adjustment protects domestic competitiveness more effectively, while an integrated emissions trading scheme achieves a greater reduction in emissions abroad. Alexeeva-Talebi et al. (2008b) conclude from their simulations of the European emissions trading scheme that market based policy measures such as the Clean Development Mechanism, allowing for flexibility in the location of emissions savings, can be effective substitutes for border tax adjustments in unilateral climate policy. Manders and Veenendaal (2008) find that border tax measures under the European emissions trading scheme significantly reduce carbon leakage. Furthermore, border tax measures appear beneficial for the EU, while they may entail a welfare loss for the rest of the world.

Finally, Lessmann et al. (2009) examine a numerical, intertemporal optimization framework with stable coalitions. They show that carbon based import tariffs increase the emissions target coalition in an welfare improving way if the tariff rate is small relative to the Armington elasticity of imports.

Our paper contributes to the CGE model based literature on border tax adjustment by assuming a carbon based tariff on commodity trade from a region without a binding emissions target to a region with a binding emissions target. Different to the literature, our paper neglects competitiveness aspects by not assuming carbon based subsidies on exports from a region with a binding emissions target to a region without a binding

emissions target.³

The first contribution of our paper is to calculate and illustrate implicit carbon contents of commodities traded between China, the industrialized countries and the developing countries based on the new GTAP 7 data for 2004 (section 3). The second contribution is to examine the effects of imposing a carbon content based border tax under a contraction and convergence climate regime with emissions trading on welfare and emissions in a stylized CGE model (section 4). Based on the results, the paper derives implications for post-Kyoto policies (section 5). The Appendix provides a description of the key parameters, variables and equations of the model. The paper starts with an overview of the underlying three region model (section 2).

2 The three region model

The underlying DART⁴ model is a recursive dynamic multi-region, multi-sector CGE model of the world economy. The static part of the model is currently calibrated to the GTAP 7 database (Narayanan and Walmsley 2008) that covers global production and trade data for countries and regions, commodities and primary factors for the benchmark year 2004. Emissions data for GTAP 7 are taken from Lee (2008). The model runs under GAMS MPS/GE. For a detailed description see Klepper and Springer (2000), Springer (2002) and Klepper et al. (2003).⁵

The version of the model scrutinized here distinguishes three regions: China (CHI), industrialized countries (IND) and developing countries (DEV). The industrialized region encompasses the OECD countries plus Hong Kong, Macao, Taiwan, Singapore and South Korea, since they are important source countries of FDI to China - and potential source countries of technology transfer to China (compare Tseng and Zebregs 2002, Whalley and Xin 2006). The model distinguishes the production factors labor, capital, land, and natural resources (fossil fuels). In order to analyze climate policies, CO₂ emissions are linked to the use of fossil fuels in production and consumption. The current sectoral aggregation covers 30 sectors in each region.

Each commodity market is perfectly competitive. Product and factor prices are fully flexible. The model incorporates two types of agents for each region: producers

³Meade (1974) and Grossman (1980) show under which conditions an equal border tax on all imports and a corresponding subsidy on all exports leads to a readjustment of the exchange rate without real economic effects. In our current analysis these criteria are not fulfilled.

⁴Dynamic Applied Regional Trade.

⁵The description in this section follows Hübler (2009).

(one producer per production sector and region) and consumers (one private and one public consumer per region). Producer behavior is derived from cost minimization for a given output. Consumers receive all income generated by providing primary factors to production processes. Consumers save a fixed share of income and invest it into capital for production in each period. Herein, investments are produced like commodities by using production inputs. The disposable income (net of savings and taxes) is then used for utility maximization by purchasing and consuming commodities. The expenditure function is modeled as a CES (constant elasticity of substitution) composite, which combines an energy bundle with a non-energy bundle.

Factor markets are perfectly competitive with full employment of all factors. Labor is a homogenous good, being mobile across industries within regions, but being internationally immobile. While in the basic version of the DART model capital is also internationally immobile, in this version capital is internationally mobile between the industrialized region and China. The benchmark values of foreign capital located in China are taken from the China Statistical Yearbook (2006, 2007). All regions are linked by bilateral trade flows, and all commodities except the investment good are traded among regions. Domestic and foreign commodities imported from different regions are imperfect (Armington) substitutes.

The model is recursive-dynamic; it solves for a sequence of static one-period equilibria for future time periods. The major exogenous, regionally different driving factors of the model dynamics are population growth, total factor productivity growth, human capital growth and investment in capital. The model assumes constant, but regionally different growth rates of human capital (educational attainment) taken from Hall and Jones (1999). Population growth rates and labor participation rates are taken from the PHOENIX model (Hilderink 2000). The resulting GDP growth paths are in line with recent projections by OECD (2008).

Technological progress has an exogenous part in every region. It consists of improvements in total factor productivity and in energy biased technological progress. In the latter case, a given output quantity can *ceteris paribus* be produced with a smaller volume of energy inputs. In China, technological progress in a certain sector additionally increases with the import intensity of the related product, with the foreign capital intensity in this sector and with forward and backward linkages across sectors within the production chain. Technological progress decreases the closer the Chinese technology

level comes to the technology frontier given by the industrialized region.⁶ This results in a process of technological convergence.⁷

3 Carbon content of trade

We calculate the implicit carbon contents of traded commodities using the GTAP 7 data set for 2004 (Narayanan and Walmsley 2008) in combination with emissions data computed from the GTAP 7 data set (Lee 2008).⁸ Such implicit carbon contents capture all emissions that occur during the production processes of commodities. Our calculation improves on Pan et al. (2008) by using the new GTAP 7 data and by distinguishing intermediate good inputs by country of origin (for detailed explanations see Ackerman et al. 2007). The latter aspect seems important for computing Chinese carbon contents of trade, since a substantial part of Chinese exports is produced by using imported intermediate goods (so that the value added is relatively low).

In the first step, we derive an input-output table, in other words a 90×90 Leontief technology matrix Λ , from the GTAP 7 data. In each column, it describes the production of a commodity i (in a sector i) in region r . The first columns contain all commodities i produced in the first region, the following columns contain all commodities produced in the second region and so on. Within each column, commodities i are listed in the same order representing the intermediate good inputs that are necessary to produce one output unit of commodity i in region r . At this point, the GTAP 7 data set does not provide *bilateral* trade flows of *intermediate* goods. It does, however, provide bilateral data on total trade flows μ (for intermediate input use plus consumption) and it does provide *bisectoral* data on total imported intermediate inputs ι of firms (without distinguishing by source country). Therefore, we use the following weighting algorithm to compute bilateral intermediate good flows ι^b from sector ii in region rr to sector i in region r :

$$\iota^b(rr, r, ii, i) = \iota(r, ii, i) \frac{\mu(rr, r, i)}{\sum^{rr} \mu(rr, r, i)} \quad (1)$$

The underlying assumption is that the distribution of source countries of imports is the same for intermediate good imports as for total imports.

In the second step, we compute the Leontief inverse χ containing the volumes of all

⁶For further details see Hübler (2009).

⁷Full technological catching up would be far beyond the time horizon of our analysis.

⁸For this section, we only need the GTAP 7 data set, not the CGE model itself. Like Peters and Hertwich (2008), we do not distinguish intermediate inputs by source country, since the GTAP data do not provide bilateral intermediate good flows.

commodities that are necessary to satisfy the demand for one unit of each commodity, and additionally to satisfy the need for intermediate inputs throughout all production stages. Herein, Ξ is a 90×90 identity matrix.

$$\chi = \Lambda \times \chi + \Xi \quad \Leftrightarrow \quad \chi = [\Xi - \Lambda]^{-1} \quad (2)$$

In the third step, we derive the direct emissions per unit of output i , denoted by the 1×90 vector ε . These direct emissions occur in each production stage via direct inputs of fossil fuels (coal, gas and oil).⁹ For this purpose, we use the data on direct emissions that Lee (2008) computes from the GTAP 7 data. She takes into account that, depending on the region, a certain share of oil and gas goes into plastic products within the chemical sector. She also takes into account that the oil sector encompasses processes where oil inputs are not burned, but refined in order to gain improved oil products. In this case, she assumes that the resulting emissions are zero.

In the fourth step, we multiply ε with χ . As a result, we obtain the 1×90 carbon intensity vector ζ that contains the emissions over all intermediate production stages that occur when producing one unit of each commodity i in each region r .

$$\zeta = \varepsilon \times \chi \quad (3)$$

Figure 1 shows the results for the benchmark year 2004.¹⁰ The figure illustrates that Chinese (CHI) products have the highest carbon (CO₂) intensities (except transportation trn), on average about 3.1kg/US\$. Especially, the Chinese carbon content of electricity generation (egw) is extremely high due to the importance of inefficient coal power in China.¹¹ As expected, commodities produced in the developing countries (DEV) have the second highest carbon contents, on average about 1.6kg/US\$, and commodities produced in the industrialized countries (IND) have the lowest carbon intensities, on average about 0.7kg/US\$.

⁹Assume, steel production uses electricity and burns oil when running machines. Then, only these direct emissions from burning oil are included at this stage of the calculation.

¹⁰We distinguish 30 sectors: agriculture and food (agr), textiles, apparel and leather (tex), beverages and tobacco (bev), business services (bui), chemicals, rubber and plastic (crp), culture and recreation (cus), coal (col), communication (com), construction (con), crude oil (cru), electricity supply (egw), electrical equipment (elm), ferrous metals (fem), financial intermediation (fin), gas (gas), machinery (mac), metal products (met), minerals (min), non-ferrous metals (nfm), non-metallic mineral products (nmm), other manufacturing (otm), paper products and publishing (pap), petroleum and coal (oil), trade and wholesale (trd), public services (pub), real estate (ree), transport machinery (trm), transportation (trn), water supply (wat), wood (woo).

¹¹The emissions intensity of gas in China was obviously an outlier. Therefore, we assumed it is equal to the emissions intensity of gas in developing countries. For further comments on accounting problems

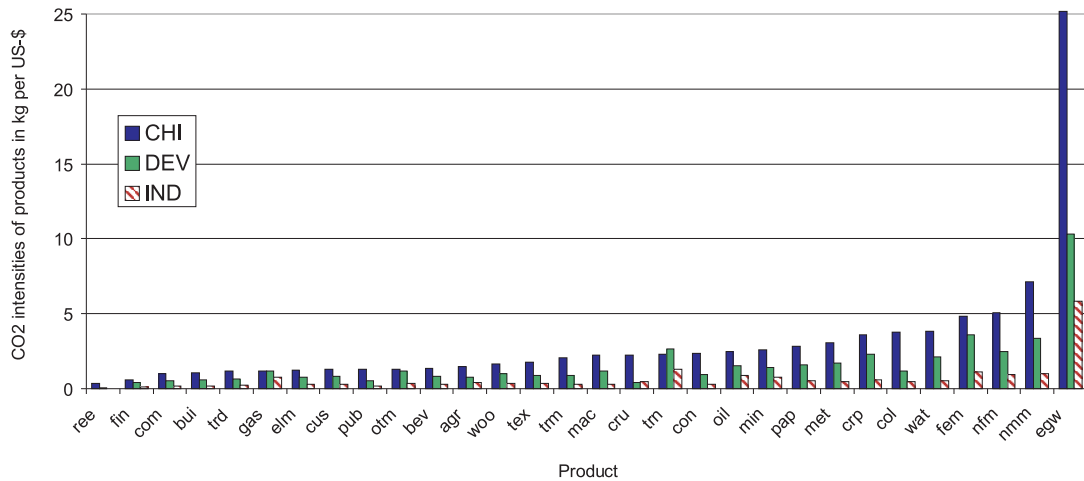


Figure 1: Carbon intensities of products

In the fifth step, *total* carbon contents of *traded* commodities per year can easily be computed by multiplying the carbon intensity factors shown in Figure 1 by the related volumes of commodity trade. (Note that implicit carbon trade within regions is not included.) Figure 2 shows the results for exports of each region. As expected, the ranking of implicit Chinese carbon export volumes is similar to the ranking of commodity export volumes. The three highest and almost equal carbon volumes are embodied in exports of textiles, apparel and leather (*tex*); electrical equipment (*elm*) and machinery (*mac*). Chemicals, rubber and plastic (*crp*) contribute the fourth highest carbon export volume which is lower than that the three highest volumes. All other products contribute lower carbon export volumes. The other developing countries obviously export substantial carbon volumes via transportation services (*trn*);¹² non-ferrous (*nfm*) and ferrous (*fem*) metals; via agricultural and food products (*agr*); via crude oil (*cru*); and via petroleum and coal products (*oil*).

Figure 3 illustrates the result of summing up over carbon contents of traded commodities per region for the benchmark year 2004. The triangle in Figure 3 visualizes the total quantities of CO₂ in Gt (Giga tons) that are implicitly traded between regions. While about 1.6Gt flow from the developing countries to the industrialized countries, China alone exports about 1.1Gt to the industrialized countries. CO₂ exports from the

in the GTAP data see Peters and Hertwich (2008) and their supporting information.

¹²The high volume of carbon exports via transportation services stems from the high export volume of transportation services given by the GTAP 7 data.

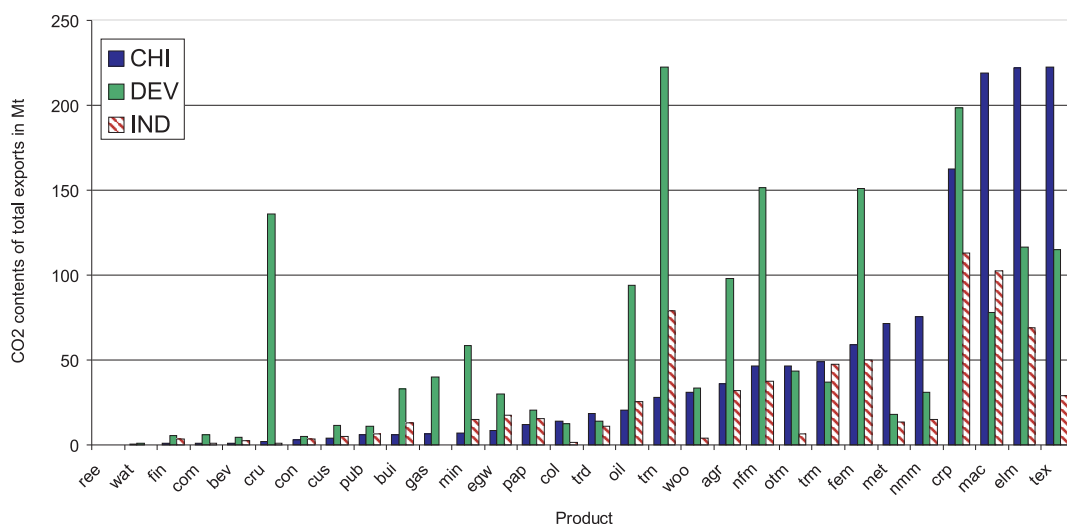


Figure 2: Implicit carbon contents of exported products

industrialized countries to the developing countries and China, as well as CO₂ flows from the developing countries to China and vice versa, are relatively low. Figure 3 does not show implicit carbon trade *within* regions. The implicit carbon trade within the industrialized region (between industrialized countries) is substantial; it amounts to 2.7Gt. The implicit carbon trade within the developing region amounts to 0.7Gt.

The percentage numbers show net CO₂ exports (implicit CO₂ exports minus imports) relative to total emissions that are actually generated in each region. As expected, China is a major net carbon exporter (24% of total Chinese emissions), while the industrialized region is a net carbon importer (15% of total emissions). The developing region is a net carbon exporter as well (12% total emissions).¹³

These outcomes indicate that a climate regime in the industrialized region alone is not sufficient. It potentially increases production in the developing countries and China and imports of the produced commodities to the industrialized region as emphasized by the carbon leakage literature. Therefore, it seems straight forward to consider policies of lowering implicit carbon trade.

¹³Compared with Peters and Hertwich (2008) who calculate the carbon contents of trade based on GTAP 6 for the year 2001, implicit carbon carbon exports of China have risen from 0.8Gt (24.4% of total Chinese emissions) in 2001 to 1.4Gt (31.3%) in 2004. Relative carbon imports of China have risen from 0.2Gt (6.6%) to 0.3Gt (7.7%). Thus, net carbon exports of China have risen from 0.6 (17.8%) to 1.1Gt (23.6% \approx 24%). According to Pan et al. (2008), China's net CO₂ exports amount to 1.7Gt in the year 2006.

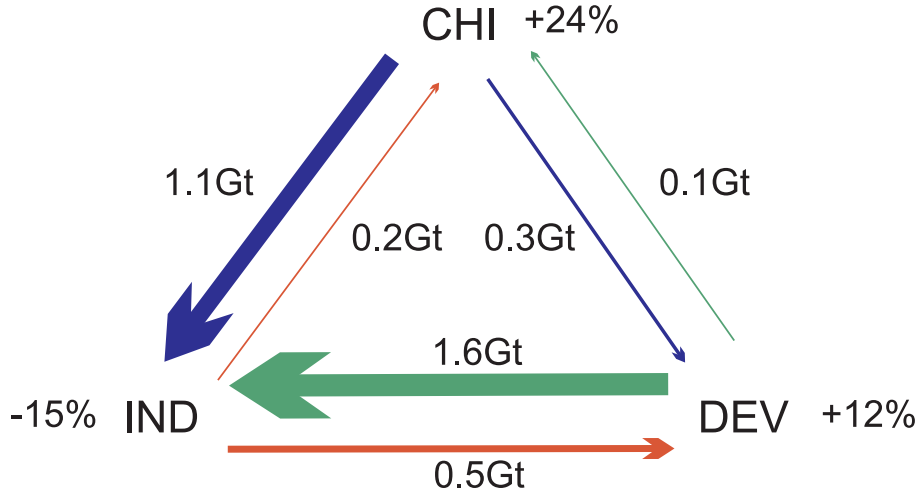


Figure 3: Interregional carbon contents of trade

4 Border tax adjustment

Thus, this section compares four policy scenarios: (1) A worldwide contraction and convergence scenario starting in 2012 including all regions, denoted by ”+chi”. Per capita emissions of the three regions converge year by year so that equal per capita emissions will be reached in the year 2050, while the model runs only until 2025 in our current analysis. In each year, regions receive emissions permits according to their current per capita based emissions goals and are allowed to trade emissions permits with the other regions. The emissions cap covers all sectors, and emissions permits can be perfectly traded across sectors and regions. (2) The same policy scenario, now excluding China, denoted by ”-chi”. This is our *reference scenario*; that is we measure accumulated welfare in other scenarios relative to this scenario. (3) The latter scenario excluding China, now with a carbon based border tax adjustment, denoted by ”-chi-bta”. The border tax revenue is received by the importing region that has a binding emissions target (IND or DEV). (4) The latter scenario with border tax adjustment, now additionally with import tariffs imposed by China as a reaction to the border tax adjustment, denoted by ”-chi-contra”. For this purpose, we assume an additional tariff rate of 5% on all products imported to China.

The regional emissions targets under the contraction and convergence regime follow the rule (Peterson and Klepper 2007):

$$\theta^{CO_2}(t, r) = \theta^{CO_2}(2011, r) \frac{2050 - t}{38} + \theta^{CO_2}(2050) \frac{t - 2012}{38}, \quad \forall t \geq 2012 \quad (4)$$

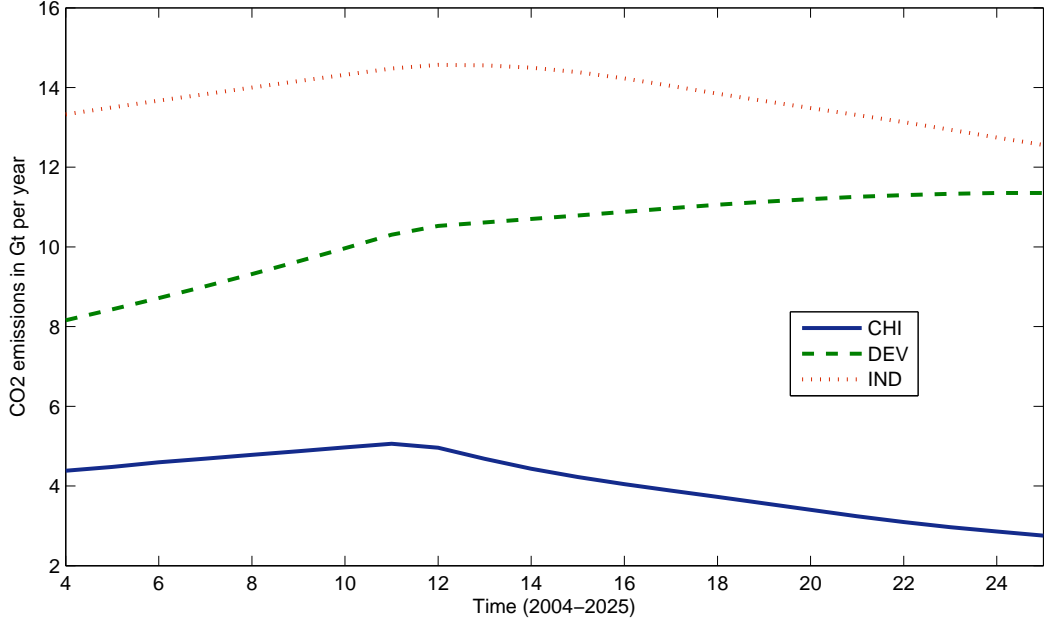


Figure 4: CO₂ emissions under a worldwide carbon price (scenario +chi)

Regional emissions in 2011, denoted by $\theta^{CO_2}(2011, r)$, are derived from the solution of the CGE for 2011. The global emissions level in 2050 is set exogenously to about 18.3Gt CO₂ which corresponds roughly to a 450ppm CO₂ intensity target (compare IPCC 2001).¹⁴ As a result, per capita emissions converge step by step from their regionally different levels in 2012 to an equalized level of 2t per capita in 2050.

The carbon based ad valorem tariff rate $\tau^{BTA}(t, CHI, r, i)$ is endogenously adjusted, where t denotes time (years), CHI denotes China as the exporting country, r denotes importing regions, and i denotes sectors or commodities. In the absence of the first best solution, a carbon price in all sectors in all regions, we aim at a second best solution by pricing imports as if they had been produced domestically. The tax rate depends on the carbon intensities of commodities that are traded from China into the industrialized or developing region, denoted by $\zeta(CHI, i)$. This implies that policy makers exactly know the real implicit carbon contents of the imported products in the benchmark year.¹⁵

¹⁴Without any climate policy, global emissions would be 39.2Gt CO₂ in 2025, and 62.7Gt CO₂ in 2050, according to our simulations.

¹⁵This is a difference to Alexeeva-Talebi et al. (2008) who assume that imported commodities are taxed as if they were produced with domestic technologies. We rather follow the scenario of integrated emissions trading, as described by Alexeeva-Talebi et al. (2008), where importers have to buy emissions permits according to the emissions that indeed occurred during the production of the imported goods. Nevertheless, we measure emissions intensities in the benchmark year and keep them constant thereafter.

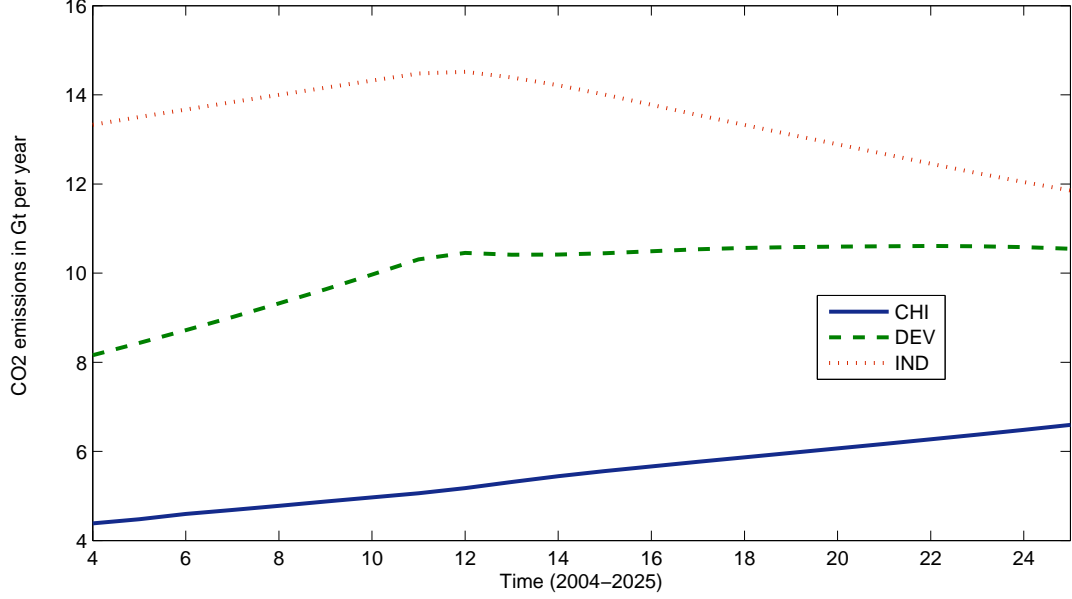


Figure 5: CO₂ emissions under a worldwide carbon price excluding China (scenario -chi)

Furthermore, the tariff rate depends on the current carbon price $p^{CO_2}(t)$ and on the current (Armington composite) import price of the commodity $p^M(i, t)$. The border tax adjustment is then given by the following constraint:¹⁶

$$\tau^{BTA}(t, CHI, r, i) = \frac{p^{CO_2}(t)}{p^M(t, r, i)} \zeta(CHI, i) \quad (5)$$

Thus, imports of commodities are due to the same carbon tax as the corresponding domestically produced commodities. As a result, across sectors, the carbon based tax rate is mainly determined by the carbon intensity. Over time, it basically follows the development of the carbon price. The carbon based tariff rate in 2012 varies between 0.06% for real estate; 0.2% for communication, public services and others; and almost 5% for gas and electricity. According to the simulation, the CO₂ price will rise up to 48 US-\$ per ton of CO₂ in 2025. As a consequence, the carbon based tariff rate will rise up to 20% for paper, oil, minerals and metals; around 30% for coal, chemicals and water; around 40% for ferrous and non-ferrous metals; more than 50% for non-metallic mineral products; and 170% for electricity, given that China's energy supply will still strongly

¹⁶Rearranging the equation and multiplying by the volume of imports $M(t, r, i)$ yields: $M(t, r, i) \cdot p^M(t, r, i) \cdot \tau^{BTA}(t, CHI, r, i) = M(t, r, i) \cdot \zeta(CHI, i) \cdot p^{CO_2}(t)$. Now, the left hand side is the total tax to be paid for importing commodity i into region r , given the ad valorem tax rate $\tau^{BTA}(t, CHI, r, i)$. The right hand side computes the carbon content of commodity i and prices it at the current carbon price.

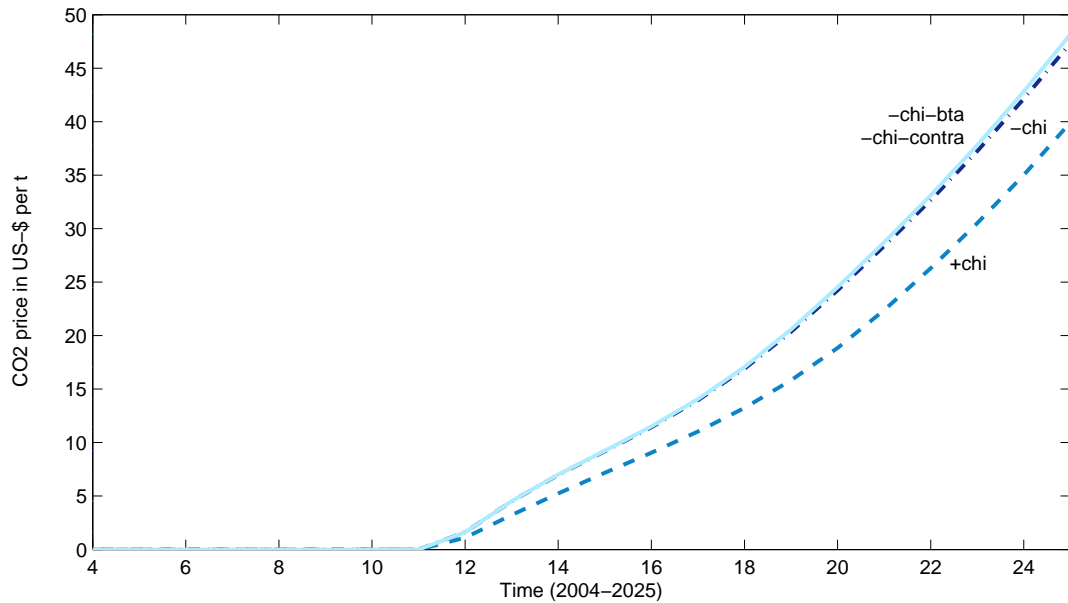


Figure 6: CO₂ price under different policy scenarios

rely on coal.¹⁷

Figure 4 illustrates the emissions paths of the three regions under a worldwide carbon price. It turns out that China will become an emissions permit seller (selling 1.5Gt of CO₂ in 2025), because it can save emissions at low marginal costs. The industrialized countries have to reduce emissions substantially over time while they buy emissions permits (amounting to 2.2Gt in 2025) in order to dampen the yearly emissions cuts. The developing countries are allowed to increase their total emissions due to their high populations and population growth. Therefore, they can achieve economic development without being hindered by tight emissions constraints. They are even willing to sell emissions permits to the industrialized countries (0.7Gt in 2025).

Figure 5 illustrates the emissions paths under a worldwide carbon price excluding China (-chi). Now, global emissions are in total higher (by almost 2Gt in 2025) than in the scenario including China (+chi). Both, the industrialized and the developing region must reduce emissions to a somewhat larger extent compared to scenario +chi, because China does no longer supply additional emissions permits, while the per capita based emissions targets for the other two regions remain as before.

This effect becomes obvious in Figure 6. When China joins the post-Kyoto regime

¹⁷Again, the gas sector in China appears as an outlier; the related border tax rate would almost be 250%.

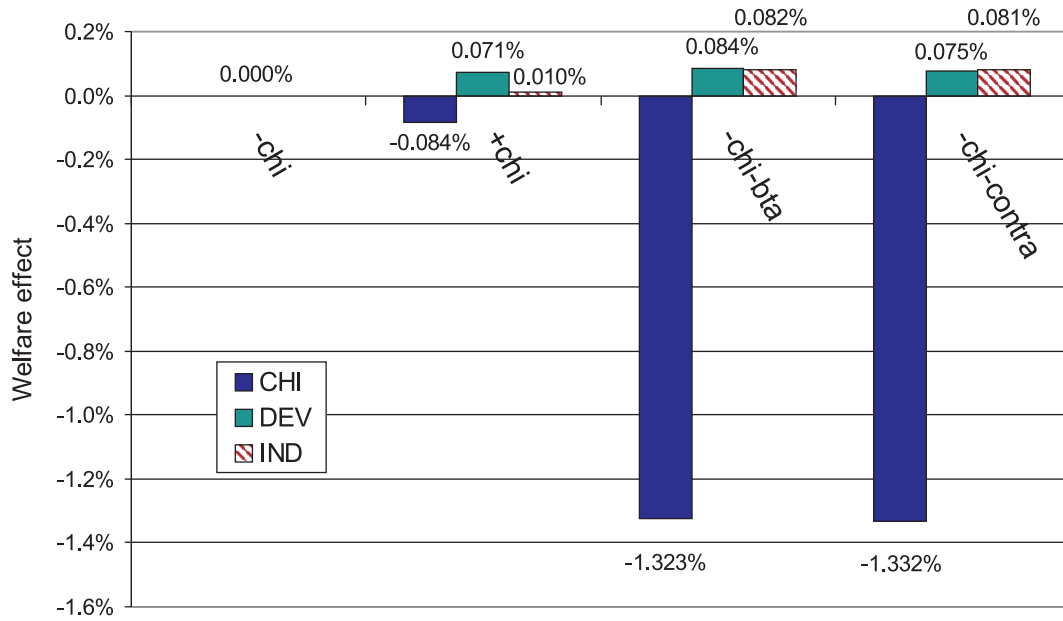


Figure 7: Accumulated, discounted welfare effects of the policy scenarios with respect to reference scenario -chi

(+chi), the time path of the carbon price is significantly lower. On the other hand, the carbon price paths almost coincide across the policy scenarios -chi, -chi-bta, and -chi-contra.

We now examine how the border tax adjustment affects emissions and exports in relation to each other. In particular, we compute the relative change in world wide emissions in scenario -chi-bta with respect to scenario -chi for each year. This relative change rises from -0.02% in 2012 to -0.72% in 2025. Additionally, we compute the relative change in Chinese exports in scenario -chi-bta with respect to scenario -chi. This relative change rises from -0.25% in 2012 to -8.64% in 2025. We then derive the following impact measure:

$$\omega(t) = \frac{EM_{-chi-bta}(t, WORLD) - EM_{-chi}(t, WORLD)}{X_{-chi-bta}(t, CHI) - X_{-chi}(t, CHI)} \quad (6)$$

This impact measure describes the change in global emissions $EM(t, WORLD)$ relative to the change in Chinese exports $X(t, CHI)$ due to the introduction of the border tax adjustment policy for each period of time t in the CGE. The CGE analysis shows that $\omega(t)$ declines from about 1.5kg/US-\$ in 2012 to about 1.0kg/US-\$ in 2025. For comparison, in section 3 we found an average carbon intensity of commodities from

China of about 3.1kg/US-\$ and of commodities from other developing countries of about 1.6kg/US-\$. The impact measure may be lower than the average carbon intensities of commodities because of replacement of part of Chinese exports by Chinese supply to the local market and by domestic supply to the local markets in the industrialized and developing region. As a consequence, emissions decline by less than the corresponding implicit carbon content of traded commodities.

Figure 7 shows accumulated, discounted welfare effects¹⁸ based on the relative Hicks equivalent variation for the different policy scenarios under scrutiny with respect to reference scenario -chi. Herein, the welfare effects do not include climate change damage, and they do not take capital stocks that remain in the final period into account. The figure reveals that the lower carbon price in scenario +chi compared to scenario -chi is significantly beneficial for the developing region. China is to a somewhat larger extent worse off than the developing region is better off, while the industrialized region is almost unaffected. The reason for this outcome is probably the higher price of Chinese exports due to the carbon tax. As a result, the industrialized region on the one hand benefits from a lower global carbon price, on the other hand suffers from higher prices of Chinese commodities which it imports to a large extent, while the developing region does not. However, these welfare effects are all rather small.

On the contrary, the introduction of a carbon based border tax under a post-Kyoto regime without China in scenario -chi-bta creates a relatively high welfare loss for China. The developing and the industrialized region benefit from the border tax revenues. On the other hand, the carbon price (as shown in Figure 6) and consequently emissions are to a very small extent affected by the border tax: Global CO₂ emissions drop by 0.1Gt in 2025, while the CO₂ price rises by 0.3 US-\$ per ton. If China reacted by imposing an additional import tariff of 5% on all commodities, all regions would be slightly worse off compared to scenario -chi-bta, while there would be practically no change in global emissions.

Finally, one side aspect is worth mentioning. In the benchmark year 2004, per capita CO₂ emissions are about 3t in China, on average 2t in the developing region, and 11t in the industrialized region. Running scenario +chi forces per capita emissions targets to converge to 2t per capita for all regions in 2050. Allowing for interregional emissions permit trading, the realized emissions in 2050 are 1t per capita in China and the developing region, and 7t per capita in the industrialized region, since the latter

¹⁸Accumulated over the time frame 2004 to 2025, discounted at a rate of 2% per year.

region strongly buys emissions permits from the former regions. Running scenario -chi, per capita emissions in 2050 would be 8t in China, 1t in the developing region, and 6t in the industrialized region. This means, the industrialized region now buys more permits from the developing countries, but in total the industrialized region buys less permits since China's emissions permits are not on the market.

5 Conclusion

We examined implicit carbon flows through commodity trade between the industrialized countries, the developing countries, and China. The large volume of carbon that is implicitly exported from China to the industrialized countries points to a substantial carbon leakage problem: If China does not join a post-Kyoto climate regime, emissions intensive production can be shifted to China, and commodities are exported to the industrialized countries, which undermines the climate regime.

Hence, border tax adjustments based on the carbon contents of traded commodities are a straightforward policy option. They can shift demand towards less carbon intensive products and locations of production. However, our analysis indicates that such policies might have small effects on the global carbon price and on global carbon emissions. They might rather make the industrialized and the other developing countries better off and China significantly worse off due to an income transfer through tax revenues. Against this background, carbon based border tax adjustment policies appear as a good menace, also as a suitable measure for collecting tax revenues from Chinese producers, but not as an appropriate measure for reducing global emissions or for reducing the global carbon price in the presence of an emissions cap.

The inclusion of China into a global post-Kyoto regime appears to be a more effective policy option with respect to the reduction of carbon emissions. China is able to save emissions substantially and to become an emissions permits seller when joining a per capita emissions based contraction and convergence regime - however at a relatively small Chinese welfare loss due to the emissions cap. The potential reason is that according to the data and the model, China has very low marginal emissions abatement costs. The additional emissions permits supply reduces the global carbon price, which creates a welfare improvement for the developing region. The industrialized region on the other hand, seems not to benefit significantly from China's inclusion into the post-Kyoto regime. The reason is probably the fact that the introduction of a carbon price in China raises the price of Chinese commodities which the industrialized region imports at a

large scale. Against this background, at first place not the industrialized countries, but the developing countries may call for an early inclusion of China into a global climate regime.

Policy makers would need to consider the following aspects. In the model, border tax adjustment is done under perfect knowledge of the implicit carbon content factors of products imported from China. In reality, true implicit carbon content factors of different products from different countries are certainly hard to estimate and to verify, which complicates the implementation of border tax adjustment policies. The opposite assumption that the carbon content factors of imported products from different countries are homogenous and equal to the carbon content factor of the corresponding domestically produced products would discriminate against exporters with low carbon intensities and benefit exporters with high carbon intensities.

On the other hand, in the current model, carbon content factors for different products are measured in the benchmark year 2004 and then kept constant. This eliminates any incentive for producers to reduce emissions in order to reduce the tax rates that are applied to their production. The opposite assumption would be that carbon content factors are truly endogenous, that is they are measured and adjusted simultaneously. Both assumptions appear not very appropriate with respect to Chinese exports. The former assumption, as it is implemented in the model, does not acknowledge efforts to reduce emissions. The latter assumption would require that Chinese firms regularly provide exact information on their energy inputs (or emissions outputs) to European or US policy makers. Perhaps, a reasonable policy would be in between both assumptions: Carbon intensity factors could be estimated (on a rough sectoral base) and updated after a certain period of time, for instance after five years. This would take energy intensity improvements of exporting economies like China into account without the necessity of large bureaucratic effort to measure and verify carbon emissions permanently. This would also create an incentive for the Chinese government to foster energy and emissions saving policies such as the Five Years Plan. Another option in between both assumptions would be the following: Policy makers estimate carbon intensity factors (on a rough sectoral base) and give firms the chance to improve the estimates by making their true emissions transparent. But herein again the problem of verifying the emissions of (Chinese) firms occurs.

However, our analysis involves numerous uncertainties, especially concerning future technological progress and economic growth. Moreover, the GTAP 7 data seem to incorporate inconsistencies between intermediate inputs in currency value terms and fossil

fuel (emissions) inputs in physical value terms and differences in accounting emissions in few cases. Therefore, the results should be treated with caution, at least in quantitative terms. The analysis rather aims at explaining potential policy outcomes qualitatively.

A detailed long-run analysis would require adjusting the carbon content factors of products over time depending on changes in the production and trade structure. Future research may also explicitly model endogenous technological progress including the rising share of renewables and possibly CCS (carbon capture and storage) since the deployment of new technologies strongly effects future emissions paths and since coal power plays a major role in China. It might turn out that international transfer of low carbon energy technologies is a more promising option than imposing trade barriers for successfully dealing with climate change.

6 References

Alexeeva-Talebi, V., A. Löschel and T. Menzel (2008a). Climate Policy and the Problem of Competitiveness: Border Tax Adjustments or Integrated Emission Trading? ZEW Discussion Paper No. 08-061, 33 pp.

Alexeeva-Talebi, V., N. Anger and A. Löschel (2008b). Alleviating Adverse Implications of EU Climate Policy on Competitiveness: The Case for Border Tax Adjustments or the Clean Development Mechanism? ZEW Discussion Paper No. 08-095, 26 pp.

Bhagwati, J. and P. C. Mavroidis (2007). Is action against US exports for failure to sign Kyoto Protocol WTO-legal? World Trade Review 6(2), 299-310.

Broder, J. M. (2009). Obama Opposes Trade Sanctions in Climate Bill. The New York Times, June, 28, 2009, online article:
http://www.nytimes.com/2009/06/29/us/politics/29climate.html?_r=1.

China Statistical Yearbook (2006, 2007). National Bureau of Statistics of China, China Statistics Press, online access:
<http://www.stats.gov.cn/english/statisticaldata/yearlydata/>.

Friedman, L. (2009). Possible Plan for Tariffs on Imports From China Remains Alive in House Climate Bill. The New York Times, June, 24, 2009, online article:
<http://www.nytimes.com/cwire/2009/06/24/24climatewire-possible-plan-for-tariffs-on-imports-from-ch-25103.html>.

Grossman, G. M. (1974). Border-Tax Adjustments - Do they distort trade? Journal of International Economics 10, 117-128.

Hall, R. E. and C. I. Jones (1999). Why do some countries produce so much more output than others? Quarterly Journal of Economics 114 (1), 83-116.

- Hilderink, H. B. M. (2000). PHOENIX plus: the population user support system version 1.0. Online information:
[http : //www.mnp.nl/en/themasites/phoenix/index.html](http://www.mnp.nl/en/themasites/phoenix/index.html).
- Hübler, M. (2009). Energy Saving Technology Diffusion via FDI and Trade: A CGE Model of China. Kiel Working Paper No. 1479, 40 pp.
- IPCC (2001). Synthesis Report - Summary for Policymakers. In: Climate Change 2001. Third Assessment Report of the Intergovernmental Panel on Climate Change, 34 pp.
- IPCC (2007). Mitigation from a cross-sectoral perspective. In: Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, UK and New York, NY, USA, 72 pp.
- Klepper, G. and K. Springer (2000). Benchmarking the Future: A Dynamic, Multi-Regional, Multi-Sectoral Trade Model for the Analysis of Climate Policies. Kiel Working Paper No. 976, 72 pp.
- Klepper, G., S. Peterson and K. Springer (2003). DART97: A Description of the Multi-regional, Multi-sectoral Trade Model for the Analysis of Climate Policies. Kiel Working Paper No. 1149, 40 pp.
- Krugman, P. (2009). The WTO is making sense. The New York Times, June, 26, 2009, online article:
[http : //krugman.blogs.nytimes.com/2009/06/26/the – wto – is – making – sense/](http://krugman.blogs.nytimes.com/2009/06/26/the-wto-is-making-sense/).
- Lee, H.-L. (2008). An Emissions Data Base for Integrated Assessment of Climate Change Policy Using GTAP. GTAP Resource #1143, GTAP CO₂ emissions (for V7), online access:
[https : //www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID = 1143](https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=1143).
- Lessmann, K., R. Marschinski and O. Edenhofer (2009). The effects of tariffs on coalition formation in a dynamic global warming game. Economic Modelling 26, 641-649.
- Liu, X. and C. Wang (2009). Quantitative analysis of CO₂ embodiment in international trade: An overview of emerging literatures. Frontiers of Environmental Science and Engineering in China 3(1), 12-19.
- Manders, T. and P. Veenendaal (2008). Border tax adjustments and the EU-ETS - A quantitative assessment. CPB Document No. 171, 36 pp.
- Marschinski, R., M. Jakob and O. Edenhofer (2009). Analysis of Carbon Leakage in an Extended Ricardo-Viner Model. Conference paper for the 2009 EAERE conference, June 24-27, Amsterdam, Netherlands, 22 pp.
- Meade, J. E. (1974). A Note on Border-Tax Adjustment. Journal of Political Economy 82(5), 1013-1015.
- Narayanan, B. G. and T. L. Walmsley (eds.) (2008). Global Trade, Assistance, and

Production: The GTAP 7 Data Base. Center for Global Trade Analysis, Purdue University, online information:

<https://www.gtap.agecon.purdue.edu/>.

OECD (2008). Environmental Outlook to 2030. Paris, France, 523 pp.

Pan, J., J. Phillips and Y. Chen (2008). China's balance of emissions embodied in trade: approaches to measurement and allocating international responsibility. *Oxford Review of Economic Policy* 24(2), 354-376.

Peters, P. and E. G. Hertwich (2008). CO₂ Embodied in International Trade with Implications for Global Climate Policy. *Environmental Science and Technology*, 42(5), 1401-1407, online access to article and to supporting information:

<http://pubs.acs.org/doi/suppl/10.1021/es072023k>.

Peterson, S. and G. Klepper (2007). Distribution Matters - Taxes vs. Emissions Trading in Post Kyoto Climate Regimes. Kiel Working Paper No. 1380, 26 pp.

Shui, B. and R. C. Harriss (2006). The Role of CO₂ Embodiment in US-China Trade. *Energy Policy* 34, 4063-4068.

Sijm, J. P. M., O. J. Kuik, M. Patel, V. Oikonomou, E. Worrell, P. Lako, E. Annevelink, G. J. Nabuurs and H. W. Elbersen (2004). Spillovers of Climate Policy - An assessment of the incidence of carbon leakage and induced technological change due to CO₂ abatement measures. Netherlands Research Programme on Climate Change Scientific Assessment and Policy Analysis, 251 pp.

Springer, K. (2002). Climate Policy in a Globalizing World. Springer, Berlin, Heidelberg, New York, 293 pp.

Tseng, W. and H. Zebregs (2002). Foreign Direct Investment in China: Some Lessons for other Countries. In: W. Tseng and M. Rodlauer (eds.), *China Competing in the Global Economy*, IMF, 68-104.

Whalley, John and X. Xin (2006). Chinas's FDI and Non-FDI Economies and the Sustainability of Future High Chinese Growth. NBER Working Paper No. 12249, 24 pp.

Wiedmann, T., M. Lenzen, K. Turner and J. Barrett (2007). Examining the Global Environmental Impact of Regional Consumption Activities, Part 2: Review of Input-Output Models for the Assessment of Environmental Impacts Embodied in Trade. *Ecological Economics* 61, 15-26.

7 Appendix

This section contains key equations of the CGE model. For further details and data sources see Klepper and Springer (2000), Springer (2002), Klepper et al. (2003) and

Hübler (2009). Tables 1 and 2 explain the meaning of the parameters and variables.¹⁹ The equations are written in quantities, while all prices are fully endogenous.

Accumulated, discounted welfare effect excluding climate change damage is derived from the relative Hicks equivalent variation of policy scenario 1 compared with reference scenario 0:

$$W(r) = \frac{\sum_{t=2004}^{2025} \{P[p^C(2004, r), U^1(t, r)] - P[p^C(2004, r), U^0(t, r)]\}(1 - \rho)^{(t-2004)}}{\sum_{t=2004}^{2025} P[p^C(2004, r), U^0(t, r)](1 - \rho)^{(t-2004)}} \quad (7)$$

Households equate expenditure to income:

$$p^C C = p^K K + p^L L + p^B B + p^{CO_2} EM + R(\cdot), \quad \forall(t, r) \quad (8)$$

Capital accumulation with a constant depreciation rate and saving rate:

$$K(t + 1, r) = [1 - \delta(r)]K(t, r) + \sigma(r)Y(t, r) \quad (9)$$

Interregional capital mobility :

$$K(t, IND) = cet[K(t, IND), F(t, CHI)] \quad (10)$$

Exogenous labor augmentation (via population growth and educational improvements):

$$L(t + 1, r) = [1 + \lambda(t, r)]L(t, r) \quad (11)$$

Basic production structure (producers minimize costs taking input and output taxes $\tau^{(\cdot)}$ into account):

$$cet(D, X) = ltf\langle N, ces\{B, cd[K, L, E]\}, \quad \forall(t, r, i), r \in \{IND, DEV\} \quad (12)$$

Basic production structure in China (producers minimize costs taking input and output taxes $\tau^{(\cdot)}$ into account):

$$cet(D, X) = ltf\langle N, ces\{B, cd[cd(K, F), L, E]\}, \quad \forall(t, CHI, i) \quad (13)$$

¹⁹The 30 production sectors are listed as a footnote in section 3.

Imported and domestically bought commodities form a consumption bundle:

$$C(t, r) = ces[D(t, r, i), M(t, r, i)] \quad (14)$$

Linking CO₂ emissions to fossil fuels (col, gas, oil) in an energy bundle:

$$E = cd\{cru, egw, ltf[EM(e), e]\}, \quad \forall(t, r) \quad (15)$$

Armington aggregation of imports from different regions (where export subsidies, and carbon and non-carbon based import tariffs $\tau^{(\cdot)}$ are imposed on traded commodities):

$$M(t, r, i) = ces\{ltf[X(t, rr, r, i), \Upsilon(rr, r, i)]\} \quad (16)$$

Exogenous total factor productivity improvement:

$$A(t+1, r, i) = [1 + \vartheta^A(r)]A(t, r, i), \quad \forall r \in \{IND, DEV\} \quad (17)$$

Exogenous and endogenous total factor productivity improvement in China:

$$A(t+1, CHI, i) = [1 + \vartheta^A(r) + T^A(t, i)]A(t, CHI, i) \quad (18)$$

Exogenous energy efficiency improvement:

$$E(t+1, r, i) = [1 - \vartheta^E(r)]E(t, r, i), \quad \forall r \in \{IND, DEV\} \quad (19)$$

Exogenous and endogenous energy efficiency improvement in China:

$$E(t+1, CHI, i) = [1 - \vartheta^E(CHI) - T^E(t, i)]E(t, CHI, i) \quad (20)$$

Herein, the strength of total factor productivity improvements in China increases with the intensities of foreign capital, of vertical linkages within the production chain, of imports, and with the distance to the technology frontier:

$$T^A(t, i) = f[FI(t, i), VI(t, i), MI(t, i)][Y_L(t, IND, i) - Y_L(t, CHI, i)] \quad (21)$$

The strength of energy efficiency improvements increases with the same factors:

$$T^E(t, i) = f[FI(t, i), VI(t, i), MI(t, i)][Y_E(t, IND, i) - Y_E(t, CHI, i)] \quad (22)$$

Symbol	Explanation
$f(\cdot)$	General function
$ces(\cdot)$ [$cet(\cdot)$]	Constant elasticity of substitution [transformation] function
$cd(\cdot)$	Cobb-Douglas function
$ltf(\cdot)$	Leontief function
t	Time, year [2004; 2025] (climate policy starts in 2012)
r [rr]	Region {IND, DEV, CHI}
i [ii]	Sector, commodity (30 sectors, see footnote in section 3)
e	Fossil fuels {col, gas, oil} (subset of i)
ρ	Time discount rate (0.02 per year)
$\delta(r)$	Capital depreciation rate
$\sigma(r)$	Saving rate
$\lambda(t, r)$	Population growth rate plus rate of educational improvement
$\vartheta^A(r)$	Rate of exogenous general technological progress
$\vartheta^E(r)$	Rate of exogenous energy biased technological progress
$\tau^{(\cdot)}(\cdot)$	Tax rate
$\Upsilon(rr, r, i)$	Transportation costs (of transporting from rr to r)

Table 1: Parameters

Symbol	Explanation
$W(r)$	Accumulated, discounted welfare effect
$U(t, r)$	Utility of the representative consumer
$P(\cdot)$	Expenditure
$C(t, r)$	Consumption (private and public)
$D(t, r, i)$	Production for domestic use
$X(t, r, i)$ [$X(t, rr, r, i)$]	Exports [bilateral trade from rr to r]
$M(t, r, i)$	Imports
$K(t, r)$ [$K(t, r, i)$]	Capital endowment [production input]
$F(t, CHI)$ [$F(t, CHI, i)$]	Endowment of CHI with capital from IND [production input]
$L(t, r)$ [$L(t, r, i)$]	Labor endowment [production input]
$B(t, r)$ [$B(t, r, i)$]	Land and natural resources endowment [production input]
$EM(t, r, e)$ [$EM(t, r, e, i)$]	CO ₂ Emissions permits endowment [production input]
$N(t, r, i)$ [$N(t, r, ii, i)$]	Intermediate good input [flow from ii to i]
$R(\cdot)$	Total tax revenue
$p(\cdot)$	Price
$A(t, r, i)$	Total factor productivity
$E(t, r, i)$	Energy input
$FI(t, i)$	Foreign capital intensity in China ($\frac{F}{K}$)
$MI(t, i)$	Import intensity in China ($\frac{M}{D+X}$)
$NI(t, ii, i)$	Intermediate good flow intensity in China (from ii to i) ($\frac{N}{D+X}$)
$VI(t, i)$	Vertical linkage intensity in China with respect to upstream u and downstream d sectors $\left[\sum_{u \neq i} FI(t, r, u) NI(t, r, u, i) + \sum_{d \neq i} FI(t, r, d) NI(t, r, i, d) \right]$
$Y_L(t, r, i)$	Labor productivity ($\frac{D+X}{L}$)
$Y_E(t, r, i)$	Energy productivity ($\frac{D+X}{E}$)
$T^A(t, i)$	Rate of endog. general tech. progress in China
$T^E(t, i)$	Rate of endog. energy biased tech. progress in China

Table 2: Variables

Chapter 5:
Costs of Climate Change:
The Effects of Rising Temperatures on Health and Productivity in Germany

Published in Ecological Economics, Elsevier, 68(1-2), 381-393
http://www.elsevier.com/wps/find/journaldescription.cws_home/503305/description#
[doi:10.1016/j.ecolecon.2008.04.010](https://doi.org/10.1016/j.ecolecon.2008.04.010)

Chapter 5:

Costs of Climate Change:

The Effects of Rising Temperatures on Health and Productivity in Germany

Michael Hübler¹, Gernot Klepper² & Sonja Peterson³

April 13, 2008

Abstract

The aim of the study is to quantify climate induced health risks for Germany. Based on high resolution climate scenarios for the period 2071 to 2100 we forecast the number of days with heat load and cold stress. The heat frequency and intensity rise overall but more in the south. Referring to empirical studies on heat induced health effects we estimate an average increase in the number of heat induced casualties by a factor of more than 3. Heat related hospitalization costs increase 6-fold not including the cost of ambulant treatment. Heat also reduces the work performance resulting in an estimated output loss of between 0.1% and 0.5% of GDP.

JEL Classifications: I10, Q51, Q54

Keywords: Costs of climate change, health effects, heat waves, mortality, hospitalization costs, labor productivity

¹ Kiel Institute for the World Economy, Düsternbrooker Weg 120, 24100 Kiel, Germany, E-mail: michael.huebler@ifw-kiel.de, Tel: +49-431-8814-401.

² Kiel Institute for the World Economy, E-mail: gernot.klepper@ifw-kiel.de, Tel: +49-431-8814-485.

³ Kiel Institute for the World Economy, E-mail: sonja.peterson@ifw-kiel.de, Tel: +49-431-8814-406.

1 Introduction

Climate change is a complex phenomenon that alters the whole environment in which humans live. Assessing the potential impacts of climate change on human health provides already a challenging task. Evaluating these effects in terms of the economic cost that these health effects may impose on an economy is even more challenging. The range of potential health effects from climate change is quite large encompassing direct effects such as the impact of heat load on human health and well-being but also indirect effects that result from climate induced storms and floods, tick-borne and food-borne diseases and allergies causing plants. Since it is difficult to attribute future economic costs to these indirect impacts, this study concentrates on the direct effects of heat load.

In the summer of 2003 thousands of people died in Germany and other European countries due to long periods of intensive heat. Yet, fatal outcomes are just the peak of a variety of heat related health risks and negative effects for human well-being and performance. Table 1 summarizes estimates of the tremendous health impacts of the heat wave 2003 – in terms of increased mortality and increased emergency hospital admissions for different European States.

There is no mono-causal relationship between temperature and detrimental health effects, though. The most important variables influencing the risk of detrimental health effects are low as well as high temperatures, humidity, wind and short- and long-wave radiant fluxes. Furthermore, different risks such as adverse physical conditions (high blood pressure, heart, kidney, liver or metabolic diseases etc.) and low physical fitness influence the effect of heat load. The main individual risk, however, is *age*. Older people (as well as young children) are most susceptible to heat, because weakness and diseases occur increasingly with higher age while the human adaptation capacity decreases. Therefore it is not surprising that fatalities due to heat load mostly occur in hospitals and nursing homes.⁴

Generally, humans are able to *adapt* to changing climatic conditions via more efficient sweating and improved blood- and fluid-circulation. We call this *physiological* adaptation. High temperatures in the first half of a year, when the affected people have not yet adapted are thus especially dangerous.⁵ Furthermore, people can adapt their *behaviour* to climate change, generally speaking by living

⁴ Calado et al. (2005).

healthier. Action plans can be prepared in hospitals and old people's homes to organize the adaptation measures during heat waves.

Place	Number of cases of heat mortality	Number of hospital emergency admissions	Source
Europe	25 000 – 35 000		Cited in Koppe et al. (2003), compare with Kosatsky (2005)
Germany	7 000		Zebisch et al. (2005)
Baden-Württemberg	1 100; 16 - 24 % increase		Cited in Koppe et al. (2003)
England	2 091; 17 % increase, 23 % increase among people aged 75 years or older, 85 % of victims older than 75 years	1 % increase among people up to 64 years, 6 % increase among people aged 75 years or older	Johnson et al. (2005)
London	616; 42 % increase, 59 % increase among people aged 75 or older	4 % increase among people up to 64 years, 16 % increase among people aged 75 or older	Johnson et al. (2005)
France	14 800; 16 % increase, 80 % of victims older than 75 years, 70 % mortality increase among people aged 75 to 94 years, 120 % mortality increase for people older than 94		EEA (2004), Kosatsky (2005)
Netherlands	650		WHO Europe (2005)
Switzerland	975; 6.9 % increase		WHO Europe (2005)
Italy	9 704, 92 % of victims older than 75 years		WHO Europe (2005), Conti et al. (2005)
Portugal	1 854; 40 % increase, 58 % up to 96.6 % of victims older than 75 years	11.6 % increase; 27.2 % increase among people aged 75 or older	Calado et al. (2005), Kovats and Jendritzky (2006)

Table 1: Estimated impacts of the European heat wave in 2003⁶

In this context, heat warning systems can help to adapt behaviour on time. They lead to the category of *technical* solutions (financed by public or private investment).

⁵ Kalkstein and Davis (1989).

Climate related building design and air-conditioning in buildings are typical technical solutions. Adaptation possibilities are related to people's social status, because poverty reduces the possibilities for heat protection through technical and structural measures, care and services. Single older people miss support and surveillance, and restricted mobility reduces the possibilities to "escape" from high temperatures (see Basu and Samet, 2002).

Against this background, the aim of this study is to quantify climate induced health risks for Germany. It addresses scientists, decision makers, medical care personnel and the public, providing findings on future health risks, so that appropriate mitigation and adaptation measures can be derived. This interdisciplinary work combining scientific knowledge from the meteorological, geographical, medical and economic field is probably one of the first attempts to systematically quantify *future* negative health effects of climate change in Germany. The focus is on such effects for which there is at least some quantitative information. For this reason we estimate *heat induced mortality, hospitalization costs and losses in labor productivity*. For mortality we also consider the effects of less coldness in winter. The estimates predict a regionally differentiated increase in heat load within Germany and consequently substantially higher negative health impacts and production losses in some regions. The mortality rise during the summer clearly dominates the possible mortality decrease during the winter. (For more detailed results than presented in this study, see Hübler and Klepper 2007). Since there are many uncertainties in past parameter values and unknown future development paths, the study is only a first step in this direction. And it has the drawbacks of a typical partial equilibrium model. The study does neither account for future changes in the sectoral pattern of the economy nor for changes of prices and quantities that are relevant for assessing future economic impact. Moreover, it does not include future adaptation, which can significantly reduce the negative effects. (The estimations are based on a population that is more or less adapted to the present climate.)

The paper proceeds as follows. In section 2 we explain forecasts of the climate model REMO⁷ based on IPCC scenarios for Germany in the period 2071-2100. Using these climate data we compute additional days p. a. (per annum) with heat load. In

⁶ For further literature reviews on heat related mortality see Basu and Samet (2002) and Kovats and Jendritzky (2006).

⁷ REMO has been criticized concerning the estimation of precipitation (the use of different time scales and deviations of the results depending on the spatial resolution). The problem has been explained and corrected (Max-Planck-Institut 2006), and precipitation is not the focus of our examination.

section 3 we derive approximations of the resulting increase in mortality as an indicator for the future health risk. In section 4 we try to estimate the economic costs of non fatal heat risks focussing on the costs of hospitalization and the reduction in labor output. In both sections 3 and 4 we apply parameter values from existing empirical studies for our estimations. Section 5 concludes.

2 Heat scenarios for Germany

Climate change causes worldwide higher temperatures with different regional patterns. To generate heat and coldness scenarios for Germany in the necessary high spatial resolution we use the Regional Climate Model REMO. The model, the underlying emission scenarios and the resulting forecasts are described in the following subsections.

2.1 Employed climate models and climate scenarios

The Regional Climate Model REMO⁸ computes climate data with a high temporal and spatial resolution (10 km times 10 km, 121 squares in the horizontal, 103 in the vertical axis) for Germany and the surroundings.

Future emissions paths depend on the uncertain development of the world economy. The forecasts therefore use the emission scenarios of the IPCC (2001) that are based on different plausible assumptions on important determinants for emissions such as economic activities and economic integration. Scenario A2 is the business as usual case. In scenarios A1B and B1 emissions rise till 2050 and then fall until 2100. While B1 implies rigorous climate protection measures, this study mainly refers to scenario A1B with a medium emissions increase.

The climate data computed by REMO are then used as inputs for the so called climate “Michel” model developed by the German Weather Service.⁹ This complete human temperature exchange model combines data on temperature, humidity, wind speed, radiation, clothing and physical activity of affected persons to calculate the so-

⁸ Jacob (2001); REMO is run by the Max Planck Institute (MPI) for Meteorology in cooperation with Deutsches Klimarechenzentrum (DKRZ), both located in Hamburg.

⁹ Fanger (1972), Gagge et al. (1986), VDI (1994 and 1998), Staiger et al. (1997), Jendritzky et al. (1990), Jendritzky et al. (2000). For a description of risk factors see Havenith (2005).

called perceived temperature for a typical reference person. Perceived temperature is a measure of how temperature affects human well-being. For instance, the heat wave in 2003 had such severe effects in France because humidity was higher than in other countries like Germany.

For practical applications it is useful to convert the perceived temperature into classes of thermal perception or classes of thermo-physiological stress according to Jendritzky et al. (2000). The 6 am and 12 am values of perceived temperatures are smoothed by applying half a Gaussian filter over 41 days. The longer ago a value, the lower its weight. The smoothed values are then used to define dynamic limits for the classes of perceived temperature composed of fixed (2/3 weight) and variable (1/3 weight) limits.¹⁰ This method imitates the human ability to adapt to climatic changes in the short-run. The procedure is also used in the heat warning system of the German Weather Service.¹¹

2.2 Heat scenarios

To generate heat scenarios for the period 2071-2100 we calculate the average number of days p. a. with heat stress from the temperature data sorted in classes of perceived temperature according to Jendritzky et al. (2000). In the illustrations heat stress encompasses the classes of strong and extreme heat stress, i. e. perceived temperatures of at least 32 °C.

The numbers of additional days with heat stress for the period 2071-2100 and for the different IPCC scenarios are given by the difference to a reference run for 1971-2000 (CTL = control). The reference run uses actually measured greenhouse gas concentrations as inputs. The total number of future hot days results from adding the expected additional number of hot days to the actually observed number of hot days in the reference period 1971-2100. The following map (Figure 1a) shows the number of days with (strong and extreme) heat stress for the REMO experiment A1B (2071-2100) minus the number of hot days in the control run CTL (1971-2000) in Germany and the surroundings in 10 km times 10 km resolution.

¹⁰ Koppe and Jendritzky (2005), Koppe (2005).

¹¹ <http://www.dwd.de/de/WundK/Warnungen/Hitzewarnung/Kriterien.htm>.

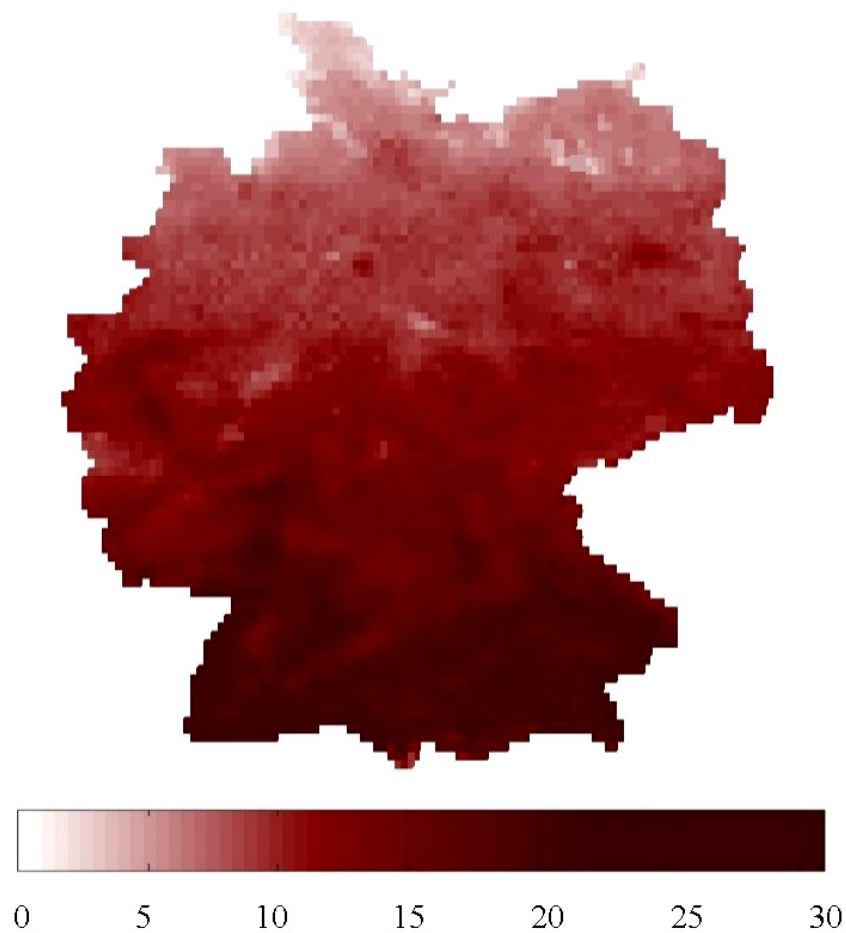


Figure 1a: Additional number of days with (strong or extreme) heat stress, REMO experiment A1B (2071-2100) minus control run CTL (1971-2000) in Germany

All experiments (B1, A1B and A2) forecast a significant increase in the frequency of days with strong or extreme heat at the end of the 21st century (2071-2100) compared to the end of the 20th century (1971-2000) and result in two to five times as many hot days. In general, the number of hot days per year rises from north to south Germany. (Note that acclimatisation will weaken the north-south gradient of heat stress.) Experiment A1B predicts one additional day with strong or extreme heat at the coast in the north, seven to 15 in the middle of Germany and 26 near the Lake Constance and in Munich (Figure 1a). Scenario B1 with lower emissions yields one additional hot day at the coast and around 18 near the Lake Constance in the south (Figure 1b). The results of scenario A2 (not shown) are very similar to those of A1B. Accordingly, the regional climatic differences will increase in Germany. Figures 1a

and 1b also reveal the heat island effect in cities like Hamburg, Berlin and Munich, that means, higher temperatures in densely populated areas compared with the surrounding areas.

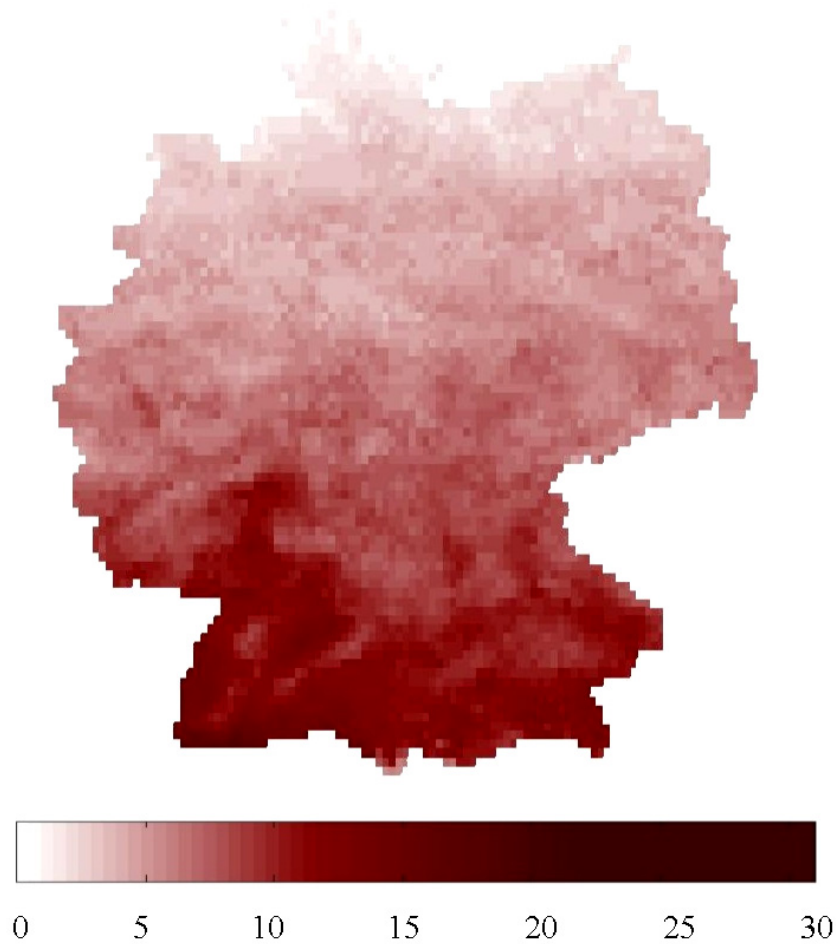


Figure 1b: Additional number of days with (strong or extreme) heat stress, REMO experiment B1 (2071-2100) minus control run CTL (1971-2000) in Germany

The comparison of Figure 1a with 1b clearly shows that a successful mitigation of greenhouse gas emissions (assumed in scenario B1) can significantly reduce future heat load.

Additionally, we use temperature time series data for the German federal states, forecasted for one city in each state like Frankfurt/Main for Hessen. Figure 2 shows for Frankfurt/Main the rising trends of heat days p. a. for B1, A1B, A2 and for the base run CTL. It is again obvious, that a successful emissions mitigation policy,

represented by scenario B1, reduces the occurrence of heat and consequently the related health risks.

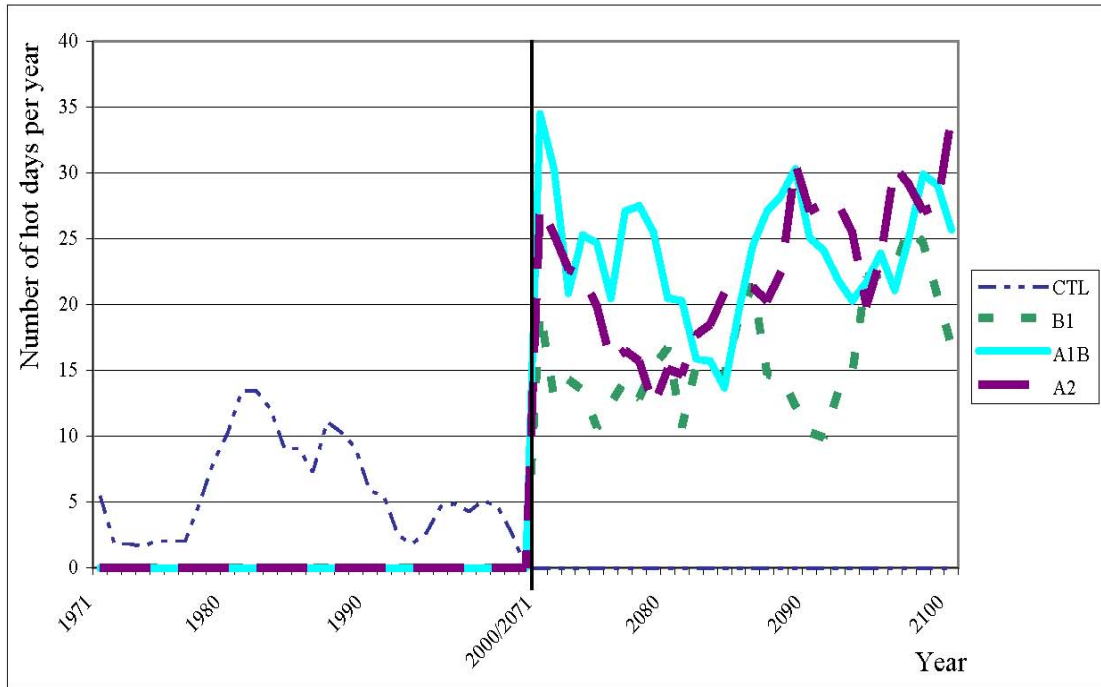


Figure 2: Additional number of days per year with strong or extreme heat stress in Frankfurt/Main (Hessen), moving averages over 5 years

2.3 Coldness scenarios

Climate change will not only increase summer temperatures and the likelihood of heat waves, but will also lead to higher average temperatures during the winter.¹²

Analogously to the heat forecast we estimate future cold stress as the average number of days per year with cold stress according to the IPCC scenarios B1, A1B and A2 for 2071-2100 and compare the results with the control run CTL for 1971-2000. We aggregate the number of days with light, modest, strong and extreme cold stress, i. e. perceived temperatures of 0 °C or lower, since there were no days with strong and extreme cold stress (perceived temperatures of -26 °C or lower) and only few days with moderate cold stress (from -13 °C to -26 °C) in the past (in CTL).

¹² It is not clear if events of strong cold stress will occur more often in the future.

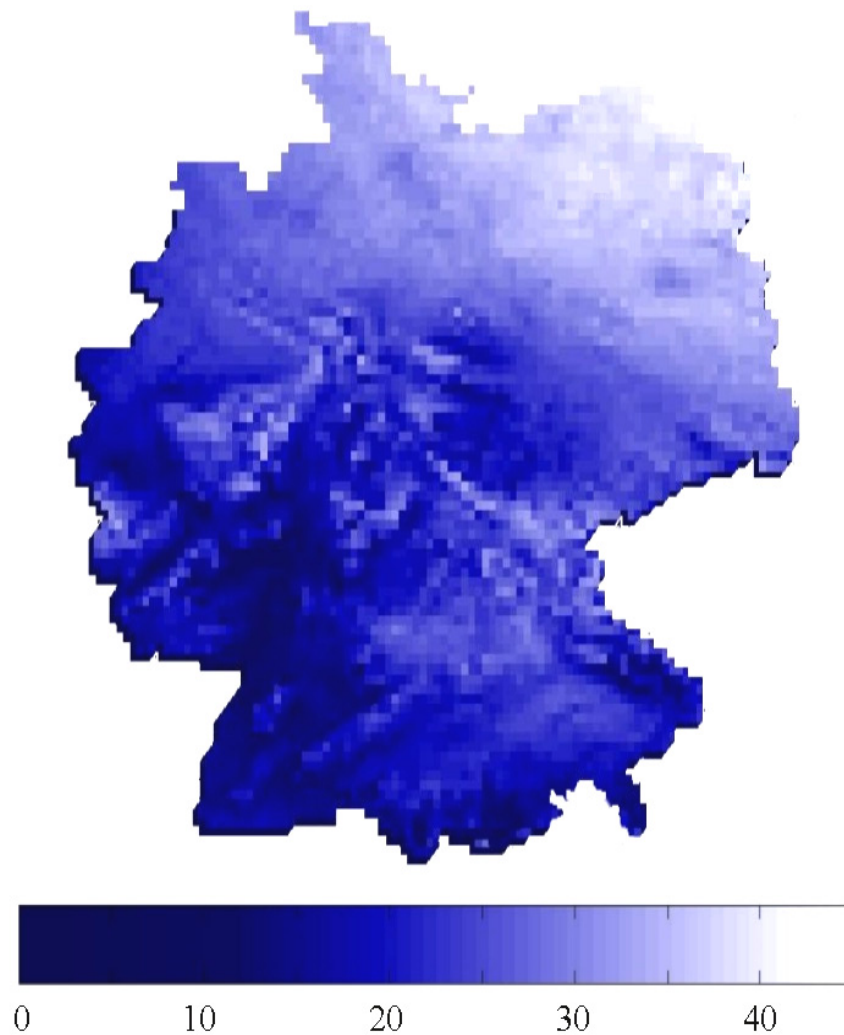


Figure 3: Reduction in the number of days per year with (light, moderate, strong or extreme) cold stress, control run CTL (1971-2000) minus REMO experiment A1B (2071-2100) in Germany

Obviously, the reduction of cold stress is highest in the north-east of Germany reaching a decrease of 30 days in scenario B1 and 44 days in A1B (Figure 3) and A2. The reduction in cold days amounts to 10 to 20 days in the middle of Germany in B1 and up to 25 days in A1B and A2. In all scenarios the lowest decrease of ca. two days per year is found at the upper Rhine rift in the south-west.¹³ Moreover, the reduction of average coldness rises with altitude, which is obvious in the higher German regions and in the Alps.

¹³ Since we include light cold stress, the reported total decrease of cold days becomes relatively high.

3 Temperature induced fatalities

In this section we use information on the relationship of temperature and mortality from the literature and combine it with the heat and coldness scenarios described before as well as with statistical population data in order to estimate future climate change induced changes in mortality. Section 3.1 refers to heat load and section 3.2 to cold stress.

3.1 Heat induced mortality

Our estimation model refers to McMichael et al. (2002) and relates heat induced mortality to the predicted additional number of hot days:¹⁴

$$D = \sum_k^K \sum_w^W (T_{w,k} + T_{0,w,k}) \cdot \frac{M_w}{365} \cdot d_k \cdot d_{season} \cdot p \quad (3.2-1)$$

D average total number of heat induced deaths p. a. in Germany 2071-2100 for IPCC scenario A1B

$T_{w,k}$ average number of additional days p. a. in 2071-2100 in perceived temperature class k in location w

$T_{0,w,k}$ actually measured average number of days p. a. in 1971-2000 in perceived temperature class k in location w

M_w absolute mortality over the whole base year 2005 in federal state w

d_k average relative mortality increase in perceived temperature class k

d_{season} seasonal mortality adjustment

p demographic change (age structure and population size) 2050 relative to 2005 in Germany

In two different runs w represents first one location for each federal state ($W = 16$) and second the 121 times 103 fields, each with a size of 10 km times 10 km. The temperature classes k ($K = 3$) are strong and extreme as well as moderate heat stress, all causing increased mortality.

¹⁴ Absolute numbers are written in capital letters, relative numbers in small letters.

$T_{w,k}$ is the average number of additional days p. a. in the prediction period (A1B, 2071-2100) compared to the reference period (CTL, 1971-2000) when the threshold of perceived temperature class k is reached in location w . $T_{w,k}$ is given by the climate forecast explained in section 2.2 and can be expressed as follows, where j ($J=30$) is the time index for years in the prediction or reference period:

$$T_{w,k} = \left[\sum_{j=1}^J (T_{j,w,k}) / J \right]_{2071-2100} - \left[\sum_{j=1}^J (T_{j,w,k}) / J \right]_{1971-2000} \quad (3.2-2)$$

In order to calculate the total number of heat related future fatalities in the first run, we add $T_{0,w,k}$, the actually measured average number of days p. a. in the reference period 1971-2000, to the estimated number of hot days. Since past observations are not available in 10 km times 10 km resolution, we cannot add $T_{0,w,k}$ in the second run.

M_w is the absolute number of deaths in general (due to any reason for death) in the base year 2005 in federal state w .¹⁵ In case of 10 km times 10 km squares, M_w is the mean mortality rate in Germany multiplied by the population size in the square.¹⁶ Dividing by 365 yields the number of deaths per day.

d_k denotes the percentage increase in general mortality M_w due to heat stress of class k . We use parameter values from Laschewski and Jendritzky (2002) for the period 1968-1997 in the federal state of Baden-Württemberg and observations during the heat wave 2003 in Baden-Württemberg reported by Koppe et al. (2003) that refer to classes of perceived temperatures. Based on Laschewski and Jendritzky (2002) we compute the mortality increase for moderate heat stress amounting to approximately 6.6 %; according to Koppe et al. (2003) the mortality increase for strong heat stress is about 9.3 %. For extreme heat stress no suitable information is available. Extrapolation yields an increase of 12.0 % (linear extrapolation) and 14.8 % (exponential extrapolation), respectively.

Furthermore, we take the seasonal adjustment of the general mortality, denoted by d_{season} , into account, because winter mortality is circa 8 % higher than the yearly average mortality while summer mortality is about 8 % lower.¹⁷ The mortality change referring to the summer is therefore multiplied by the factor 0.92.

¹⁵ Statistisches Bundesamt (2006a).

¹⁶ Data from Landscan (2001).

¹⁷ Calculation based on Laschewski and Jendritzky (2002).

p is the demographic adjustment coefficient. The number of people below the age of 75 will decrease by 18 % ($p(74-) = 0.82$) until 2050, and the number of people aged 75 years and more will increase by 95 % ($p(75+) = 1.95$) resulting in a decrease of total German population by 9.3 %.¹⁸ In accordance with the experiences of the heat wave in 2003 we assume that 80 % of all heat stress victims are people aged 75 years and more ($m(75+) = 0.8$) and carry out sensitivity analyses to control for different parameter assumptions.¹⁹ The following formula captures the demographic change:

$$p = m(75+) \cdot p(75+) + (1 - m(75+)) \cdot p(74-) \quad (3.2-3)$$

$m(75+)$ share of people of age 75 and more among heat fatalities during the heat wave 2003

$p(75+)$ number of people of age 75 and more in 2050 relative to 2005

$p(74-)$ number of people of age 74 and less in 2050 relative to 2005

Since there is a lack of more disaggregated information, we assume that the demographic development is the same across all federal states and 10 km times 10 km squares. Moreover, we neglect population movements within Germany (that will mainly take place from the east to the west and to the south of Germany) as well as international migration.

The calculations result in a substantially increased heat related mortality at the end of the 21st century. The first run on federal state level yields on average ca. 16 700 heat induced deaths p. a. in the period 2071 to 2100 (using exponential extrapolation for the class of extreme heat stress) which can be disaggregated as follows: Today's number of deaths amounts to ca. 4 500 (bottom part in Figure 4). This value is not measured but is generated by the model as a reference. Without demographic change the prediction yields about 5 200 additional heat induced deaths (middle part in Figure 4). The joint effect of a decrease in total population and an almost doubling number of elderly people creates ca. 7 000 additional potential heat victims, since the latter effect dominates (upper part in Figure 4). This means, the total number of heat induced fatalities rises by the factor 3.7. Excluding today's

¹⁸ Calculation with data from Statistisches Bundesamt (2006a). We use population forecasts for 2050 since the forecasts by Statistisches Bundesamt do not include the period 2051-2100.

¹⁹ Shares of affected elderly people reported in the literature vary between 44 % (in the USA) and 96.6 % (in Portugal); overview in table 1; EEA (2004), p. 74; Morbidity and Mortality Weekly Report for the USA cited in Uphoff and Hauri (2005).

number of heat related deaths from the calculation leads to ca. 9 000 additional future deaths p. a.

Accordingly, the future health risk increases substantially without appropriate mitigation and adaptation efforts. It is important to note that these estimations do not include any adaptation to climate change and hence overestimate the real effects. Furthermore, we do not take into account the so-called “harvesting effect”, meaning that a certain number of sick and elderly people might have died even without heat in the near future. The heat event shifts the date of death forward. As a consequence the mortality ratio can decrease under its average level after the heat event.

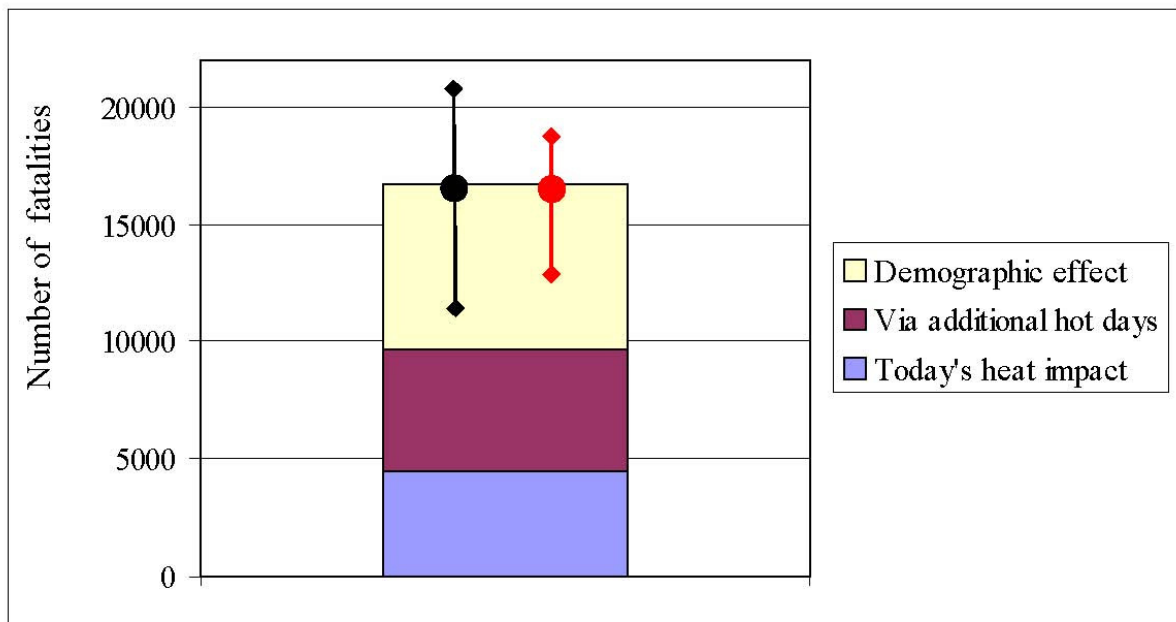


Figure 4: Number of heat induced deaths p. a. in Germany, 2071-2100 (exponential extrapolation for extreme heat stress)

Subtracting these casualties will in the short-run probably result in an up to 25 % reduction of the mortality numbers for Germany. However, Koppe (2005) finds (referring to Baden-Württemberg, 1968-2003) that the extent of the “harvesting effect” differs substantially, being much higher for people younger than 75 years (85.4 % for moderate heat load and 50.7 % for strong heat load) than for people in the age of at least 75 years (17.7 % for moderate heat load, only 2.1 % for strong heat load).

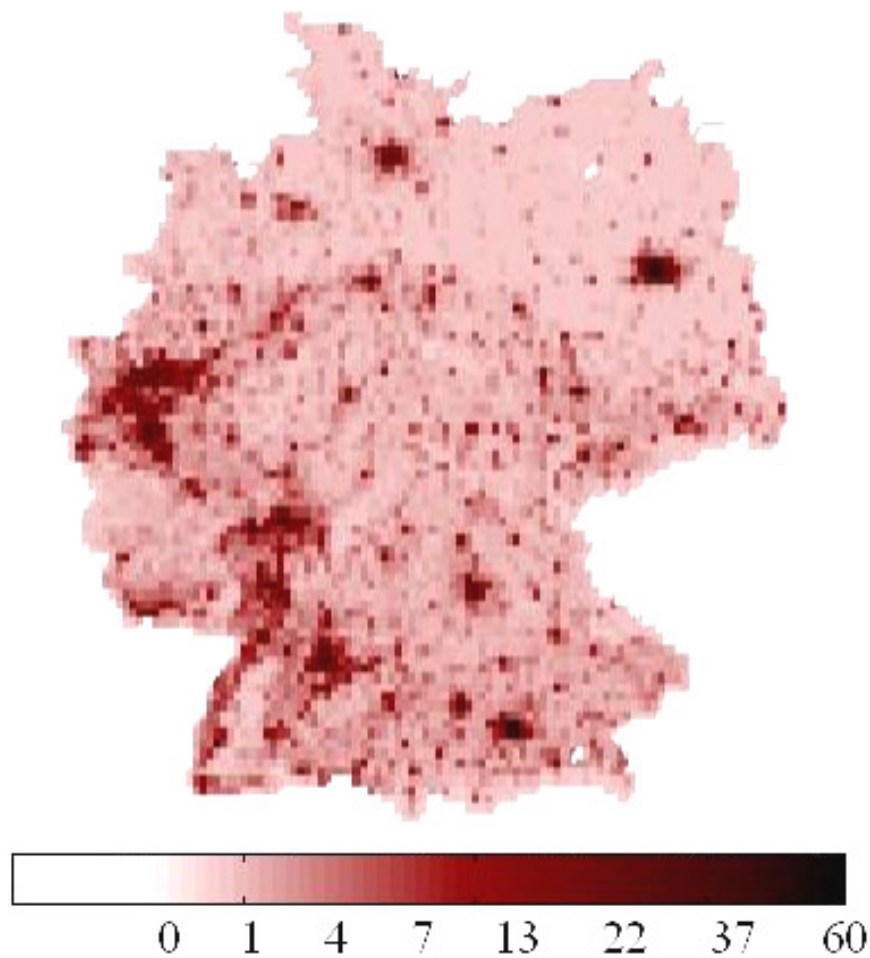


Figure 5: Distribution of additional heat induced deaths p. a. per 10 km times 10 km sectors in Germany, 2071-2100, A1B (exponential extrapolation for extreme heat stress)

Furthermore, we carry out sensitivity analyses for the parameters “heat related mortality increase d_k ” and “share of elderly people among heat victims $m(75+)$ ”. Both vary with time and region and depend on the adaptation status and adaptation ability of the affected population and can thus not be determined exactly. The vertical sensitivity bar on the left hand side shows the range of total outcomes from about 11 500 to about 21 500 when d_k varies by $\pm 30\%$ in accordance with the range of findings in the literature. The vertical sensitivity bar on the right hand side of Figure 4 shows the range of total mortality from ca. 13 500 to ca. 19 000 when $m(75+)$ is changed between 50 % and 100 %, again referring to the findings in the literature (see Table 1). In the second run we compute heat induced mortality in 10 km times 10 km

resolution visualized in Figure 5. Obviously, most heat fatalities will be in the regions with the highest population densities, i. e. in the cities, and the health risk increases from north-east to south-west (in accordance with Figure 1a).

Adding up the numbers of all 10 km times 10 km sectors yields the average additional number of heat related deaths in Germany p. a. amounting to 8 500 (compared to almost 9 000 in the first run ignoring the base value of today's observations).²⁰

These results have a similar order of magnitude as studies for the UK, Portugal and Australia.²¹ A test run using the number of hot days in the year 2003 as an input indeed reproduces the estimated ca. 7 000 heat victims in Germany.

3.2 Coldness induced mortality

While we expect dangerous health impacts in summer, there might be positive effects in winter due to less coldness. There are much less empirical studies on the relationship of health and coldness and the role of age as a risk factor than for heat. The time lags of low temperature events and health effects are much longer than in summer, and the correlation is statistically weaker. The medical causality is different, and it is not evident, whether the health risk steadily increases with lower temperatures. To get an idea of the effects in winter we nevertheless apply again model equation (3.2-1), this time in high spatial resolution only (w refers to 10 km times 10 km sectors) with the following new parameters:

$T_{w,k}$ is now the average reduction of the number of days p. a. with cold stress in the prediction period (A1B, 2071-2100) compared to the reference period (CTL, 1971-2000). k encompasses the classes of perceived temperature with light and moderate cold stress.

$T_{0,w,k}$, the actually measured average number of days p. a. with cold stress in the reference period 1971-2000, is not available in the high resolution and therefore not included here.

²⁰ Same demographic adjustment as in the first run. Ignoring the base value means excluding the lower part in figure 6 as well as part of the upper (demographic) part in figure 6, because today's base value is expanded by the demographic factor, which is included in the upper part.

²¹ Overview of heat victims prognoses in Kovats and Jendritzky (2006, p. 87)

M_w is the general mean mortality rate of Germany multiplied by the population size in a 10 km times 10 km square as before.²² Dividing by 365 in the formula yields the number of deaths per day.

d_k is the relative increase in general mortality M_w due to cold stress of class k . Since we expect less days with cold stress in the future, this mortality increase will occur less frequently. We use again parameter values from Laschewski and Jendritzky (2002) that refer to classes of perceived temperature. The value for moderate cold stress is ca. 9.3 %. The value for light cold stress is computed as the average of mortality from 0 to -12 °C amounting to approximately 5.7 %.²³

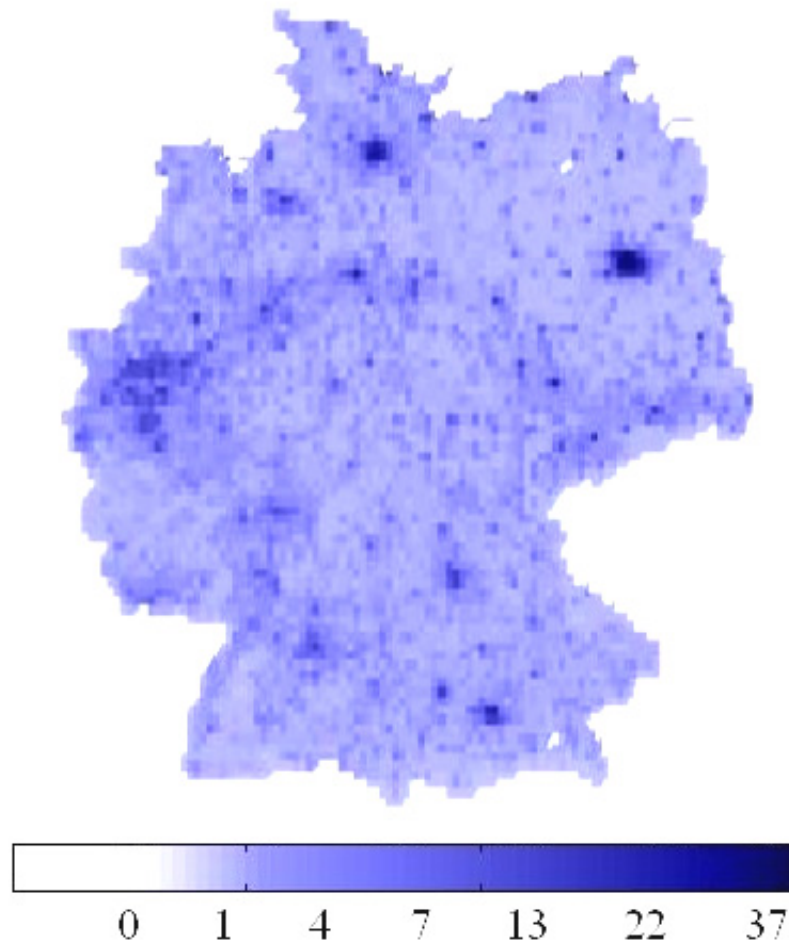


Figure 6: Distribution of the reduction in coldness induced deaths p. a. per 10 km times 10 km sectors in Germany, 2071-2100, A1B

²² Data from Landscan (2001).

²³ Laschewski and Jendritzky (2002) eliminate the impact of influenza epidemics by smoothing outliers in the mortality data.

General mortality in winter is about 8 % higher than the yearly average. Hence, the adjustment coefficient d_{season} is 1.08.

According to Hassi (2005) elderly people are very susceptible to cold stress. Thus, it appears plausible to apply the same demographic adjustment factor p as for heat load making both results comparable.²⁴ (80 % of the affected people are at least 75 years old.)

The resulting map (Figure 6) shows the highest expected reduction in winter mortality in areas with the highest population densities, i. e. in the cities. Comparing Hamburg, Berlin and Munich reveals higher mortality risk reductions in the north and north east compared to the south of Germany (in accordance with Figure 3).

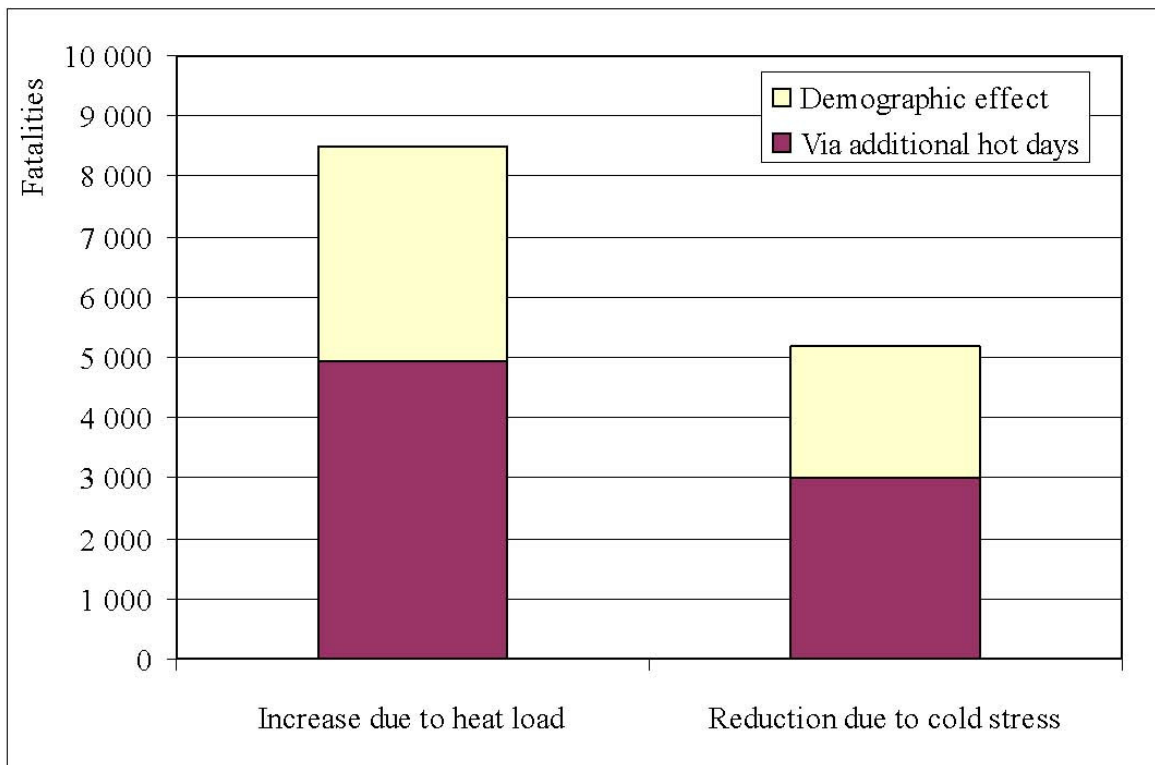


Figure 7: Comparison of the increase in heat related deaths and the reduction in coldness related deaths p. a. in Germany, 2071-2100, A1B, both computed in 10 km times 10 km resolution without base observations from the past

Adding up the numbers of all 10 km times 10 km sectors yields the average reduction in heat related deaths in Germany p. a. amounting to ca. 5 200 compared with 8 500 p. a. due to heat in summer, assuming the same demographic adjustment (age effect).

Hence, on average, the dangers of heat dominate the possible health advantages of milder winters by far, and the uncertainties of the estimations for the winter are higher than those of the summer. Figure 7 visualizes these results. A calculation without the age effect leads to only ca. 3 000 p. a.

A comparison of Figures 5 and 6 shows that in most parts of Germany the positive and negative thermal effects roughly balance each other. While only the north-eastern region with its mild maritime climate can benefit from a positive net effect, dangerous heat impacts dominate in the south and south-west. This outcome has to be interpreted with caution though: A zero or slightly positive net effect in some regions does not mean that there is no need to react. Even though there might be advantages in winter, the harmful impacts in summer can be avoided via natural adaptation and feasible adaptation activities.

4 Economic costs of heat

Casualties represent the most extreme danger of heat. As described in Table 1, the European heat wave in 2003 also caused a rise in hospital emergency admissions. This fact is confirmed by a study for the USA: Semenza et al. (1999) find 11 % more hospital emergency admissions during the heat wave 1995 in Chicago in general and 35 % more admissions in the age group 65 years and more. Nevertheless, the empirical evidence is much weaker than for heat related mortality. Moreover, heat can negatively affect well-being making us feel uncomfortable and exhausted and entails difficulties to concentrate. These aspects are hard to quantify, but become evident when doing physical or mental work.

Estimating the economic costs of the different effects of heat entails a number of methodological problems. The main problem is that our temperature scenarios are for the time period 2071 to 2100, while it is not possible to obtain resilient forecasts of the development of the German society and economy over the next 100 years. We thus assess the costs of heat referring to the current gross national income, to the current sectoral structure, to current prices and to today's (medical and general)

²⁴ Data on England show a higher coldness risk for people aged more than 75 years.

technologies. There is also a lack of hospitalization data referring to perceived temperatures. Again, we neglect any kind of adaptation to climate change.

In section 4.1 we use the indicator “hospital emergency admissions” to get an idea of the costs associated with serious illnesses. In section 4.2 we obtain a rough estimate of the effects of heat on labor productivity. (We abstain from a monetary valuation of the fatalities estimated in section 3.1.)

4.1 Heat induced hospitalization costs

Since there is no suitable information on medical treatment in practises and the related costs of treatment and medication, our estimation refers to statistics on the influence of heat on hospital emergency admissions and hospitalization costs. Moreover, it is difficult to identify the quantitative influence of temperature on specific diseases and the costs directly related to those diseases.

The projection of hospital costs is based on the estimation of additional heat days for the 16 reference places in different federal states for the IPCC scenario A1B in 2071 to 2100. The estimation model²⁵ is similar to formula (3.2-1):²⁶

$$H = \sum_k^K \sum_w^W (T_{w,k} + T_{0,w,k}) \cdot \frac{S_w}{365} \cdot K_w \cdot h_k^p \quad (4.2-1)$$

H average total hospitalization costs p. a. in Germany in the prediction period 2071-2100 for IPCC scenario A1B

$T_{w,k}$ average number of additional days p. a. in the prediction period in perceived temperature class k in location (federal state) w

$T_{0,w,k}$ actually measured average number of days p. a. in the reference period 1971-2000 in perceived temperature class k in location w

S_w absolute number of new patients over the whole base year 2004 in federal state w

K_w general hospitalization costs per case in federal state w

²⁵ Similar McMichael et al. (2002) predicting mortality increases.

²⁶ Absolute numbers are written in capital letters, relative numbers in small letters.

h_k^p average relative increase in the number of hospital emergency admissions in perceived temperature class k , the demographic change p (age structure and population size) 2050 relative to 2005 in Germany is included

$T_{w,k}$ and $T_{0,w,k}$ are given by the climate data as before.

S_w denotes the base number of new patients over the whole base year 2004 in federal state w and K_w the related average medical treatment costs per case of medical treatment.²⁷

h_k^p is the average relative increase in the number of hospital emergency admissions in the classes of moderate, strong and extreme heat load including the demographic development. There is no information available on the relationship of perceived temperature and emergency cases in Germany. Thus, we apply parameter values in absolute temperatures from a study for England during the heat wave 2003. In this statistical analysis Johnson et al. (2005) find a 1 % rise in hospital emergency admissions in the age group up to 64 years and a 6 % increase for people aged 75 years and more at maximal temperatures in the range of about 25 to 31 °C. According to Johnson et al. (2005) hospital emergency admissions decrease by 4 % in the age group 65 to 74 years. These values are applied to the class of moderate heat load referring to the temperature range during the heat wave 2003 in England. In London maximum temperatures during the heat wave 2003 were in the range of 35 to 38 °C. So, we apply the following numbers to the class of strong heat stress: A 4 % increase of hospital admissions in the age class up to 64 years, a 5 % decrease in the age group 65 to 74 years and a 16 % rise among people aged 75 and older. Due to a lack of information on extreme heat stress we compute the related hospital admission changes via (linear and exponential) extrapolation.

The relative mean change h_k^p can then be expressed in the following way:

$$h_k^p = share(64-) \cdot p(64-) \cdot h_k(64-) + share(65-74) \cdot p(65-74) \cdot h_k(65-74) \\ + share(75+) \cdot p(75+) \cdot h_k(75+) \quad (4.2-2)$$

Where $share(.)$ is the share of people in the age group in parentheses among the people who left hospital in the reference year 2004.²⁸

²⁷ Statistisches Bundesamt (2004a), Statistisches Bundesamt (2004b).

²⁸ Statistisches Bundesamt (2006b).

$p(.)$ denotes the size of the age group in parentheses in 2050 relative to 2004.²⁹ (This includes the change of the whole population size.) We assume that the demographic development is the same across all federal states.

$h_k(.)$ are the parameter values for the classes k of moderate, strong and extreme heat stress and the age groups in parentheses derived from England and London as described above.

A methodologically precise calculation requires the following consideration: Hospitalization costs in the reference year 2004 already include heat induced costs. So, before calculating the future heat related cost increase, today's heat induced costs need to be subtracted:

$$H_0 = h_k^p \cdot (S_w \cdot K_w - H_0) \Leftrightarrow H_0 = \frac{h_k^p}{1 + h_k^p} \cdot S_w \cdot K_w \quad (4.2-3)$$

H_0 denotes heat related hospitalization costs in the reference year. Thus, total reference year hospitalization costs net of heat effects become slightly smaller than before, because we diminish base costs $S_w \cdot K_w$ by H_0 :

$$H = \sum_k^K \sum_w^W (T_{w,k} + T_{0,w,k}) \cdot \frac{S_w \cdot K_w - H_0}{365} \cdot h_k^p \quad (4.2-4)$$

The result shown in Figure 8 indicates average hospitalization costs of about 495 million € per year in the period 2071 to 2100. This corresponds to an increase by a factor 6 compared to the climatic reference period 1971 to 2000. On the other hand, this number represents only 0.88 % of total German hospitalization costs (for all kinds of diseases) and 0.27 % of German health care expenses. Figure 8 shows that more frequent and more intensive heat directly causes costs of ca. 222 million € (middle part), while the demographic change (age effect) contributes 191 million € (upper part). The base value of 82 million € (lower part) was not measured in the past but is generated by the model.

The reaction of emergency cases to heat h_k^p is a crucial parameter. Hence, we carry out a sensitivity analysis varying the impact of heat on emergency cases by ± 30 % according to the magnitudes found in the literature. This yields total hospitalization

²⁹ Statistisches Bundesamt (2006a).

costs in the range of 300 to 700 million € p. a., represented by the vertical bar in Figure 8.

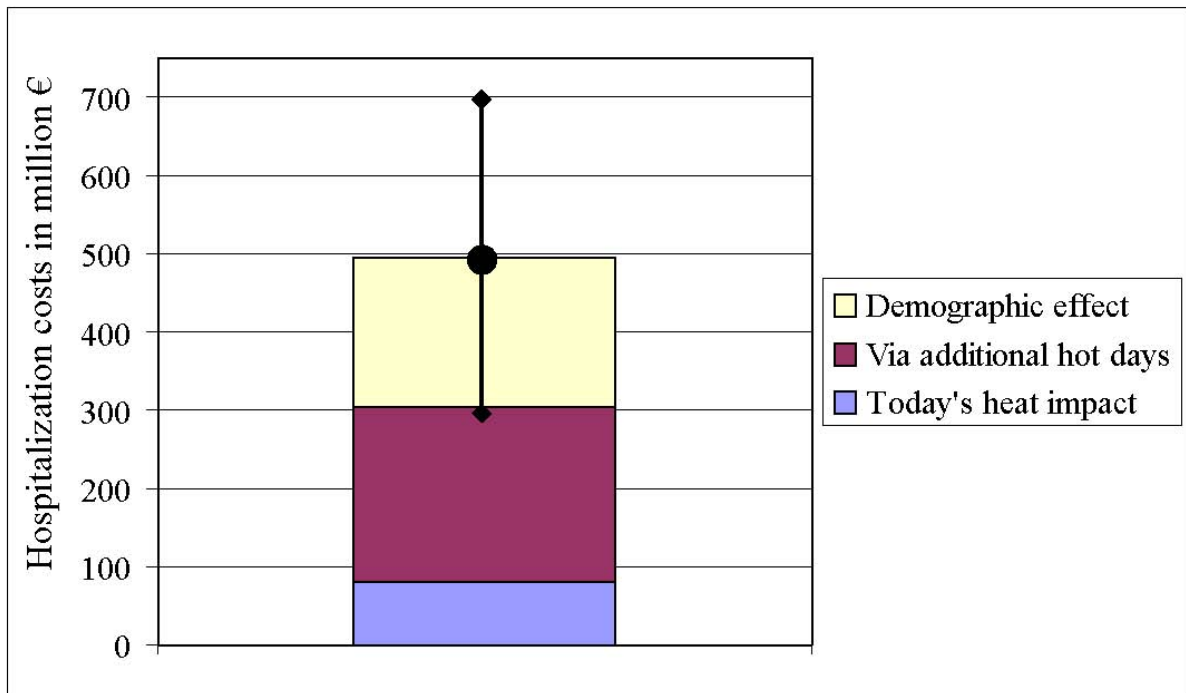


Figure 8: Hospitalization costs p.a. in Germany, 2071-2100 (exponential extrapolation for extreme heat stress)

4.2 Heat induced production loss

A number of studies investigate work performance or mental and mechanical abilities under different thermal environments and find evidence for strong negative effects of temperatures above the most comfortable level of slightly more than 20 °C. The studies describe human performance reductions in a range of 3 % to 50 % for temperatures higher than the comfortable level, reaching up to 75 % at temperatures of 35 to 37 °C.³⁰ For instance, office staff reached the maximal performance at 23 °C and only 70 % of the maximum at 30 °C. However, the range of results is large, and the economic interactions are complex. A more detailed forecast would need to distinguish between indoor and outdoor work as well as mental and physical work. These aspects have not been implemented in this first attempt and leave room for

³⁰ Wyon (1986), Kampmann (2000), Parsons (2003), Bux (2006).

further research. In this calculation we do not take into account any effects of coldness, rain or storms on production, either.

Assuming that heat directly reduces labor output, from a macroeconomic point of view the estimation model has the following form:³¹

$$L = \sum_k^K \sum_w^W T_{w,k} \cdot \frac{GDP_w + L_0}{365} \cdot q \cdot g_k \quad (4.3-1)$$

$$L_0 = g_k \cdot (L_0 + GDP_w) \Leftrightarrow L_0 = \frac{g_k}{1 - g_k} \cdot GDP_w \quad (4.3-2)$$

L average GDP loss p. a. in Germany in the prediction period 2071-2100 for IPCC scenario A1B

L_0 heat related GDP loss in Germany in the reference year 2004 (generated by the model)

$T_{w,k}$ average number of additional days p. a. in the prediction period in perceived temperature class k in location (federal state) w

GDP_w gross domestic product in the reference year 2004 in federal state w

q wage share in Germany in the reference year 2004

g_k mean relative productivity reduction when the threshold of perceived temperature class k is reached

GDP_w is the gross domestic product on German federal state level in 2004. This implies that the forecast refers to the GDP in the reference year 2004, because the development of GDP cannot be predicted till the end of the 21st century. Demographic change is neglected as well as technological progress, since both aspects cannot properly be implemented in a simple model. A possible interpretation is that the output expansion via technological progress just compensates the output loss due to a shrinking and ageing society.

As in the calculation of hospital costs we account for *today's* negative heat effects. German GDP would be higher, if temperatures were in the range of slightly more than 20 °C that is most comfortable for human beings, during the whole year. Equations (4.3-2) describe how to compute today's heat impact on production L_0 .

³¹ Absolute numbers are written in capital letters, relative numbers in small letters.

The wage share q is defined as labor income relative to total income in the economy, which is 68.4 % in Germany in the reference year 2004. We assume the same wage share in all federal states.³² It is unclear to what extent the productivity of machines, controlled by people who suffer from heat, also decreases. Thus, production losses focus on the production factor labor only. Again, changes of q due to the demographic and technological progress cannot be predicted in this simple model.

Since the quantitative impact of heat on work performance is unclear, we make the conservative assumption that there is no negative effect of moderate heat load and apply the range of scientific results cited by Bux (2006) for g_k , i. e. a productivity reduction of 3 % to 12 %.

Figure 9 shows the economic loss due to a heat induced labor productivity decrease under two different impact assumptions. In the first case labor productivity is assumed to fall by 3 % on days with strong or extreme heat stress. Today's reference heat loss generated by the model amounts to approximately 540 million € (lower part), equal to 0.03 % of today's GDP; the estimated future loss is almost 2 billion € (upper part), together ca. 2.5 billion € or 0.12 % of today's GDP.

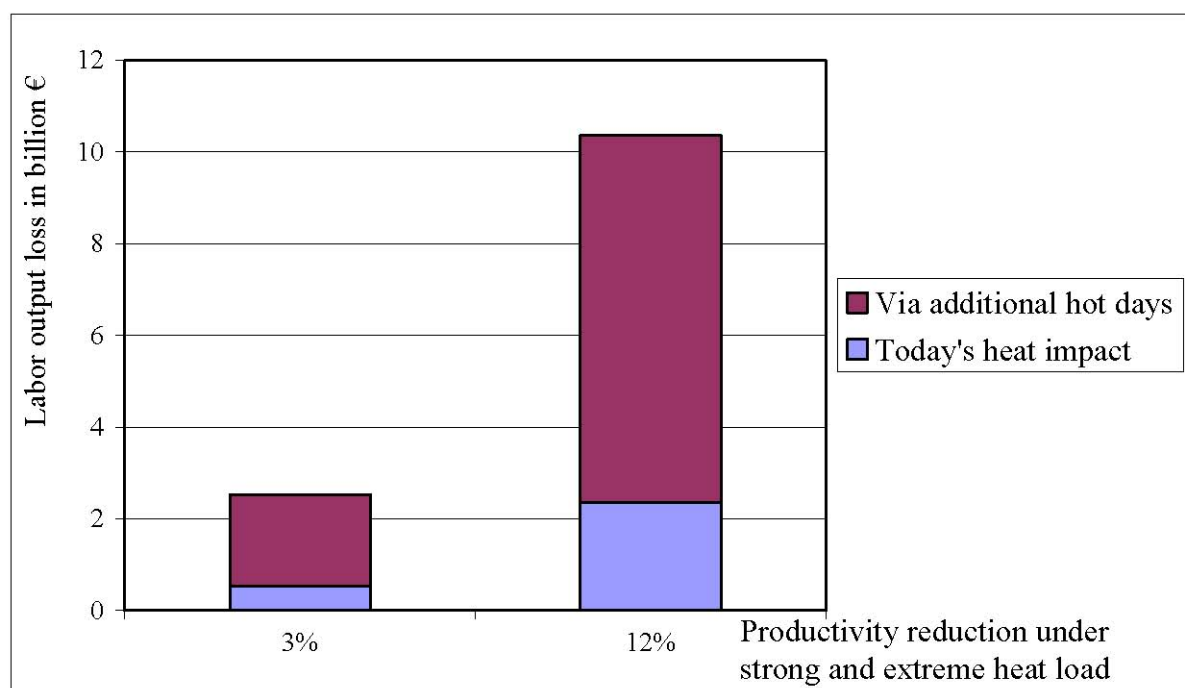


Figure 9: Heat related production loss in Germany p. a., 2071-2100 under two different impact assumptions

³² Statistisches Bundesamt (2006c), Volkswirtschaftliche Gesamtrechnung, 24.4, referring to 2004.

Assuming g_k equal to 12 % for strong and extreme heat, the estimation results in a base value of about 2.4 billion € or 0.11 % of today's GDP and in an additional future heat loss of almost 8 billion €, in total ca. 10.4 billion € or 0.48 % of today's GDP.

If strong heat stress causes a productivity loss of 3 % and extreme heat a loss of 12 %, the resulting total economic loss will be 6 billion €, which is between the results described above.

Using numbers of 30 % to 50 % mentioned in the literature for g_k , the heat related losses today (2.7 % of today's GDP) and in the future (5 % of today's GDP) are much higher. These outcomes need to be treated with caution though, since this first attempt does not differentiate between different kinds of work.

The calculations are based on IPCC scenario A1B. Using IPCC scenario B1 (low emissions) and a 12 % heat impact on labor productivity yields an additional loss of ca. 4.2 billion €, which is significantly lower than in the A1B scenario (almost 8 billion €, representing the expected emissions development). Consequently, a successful reduction of greenhouse gas emissions clearly lowers the economic loss.

Figure 10 illustrates the different magnitudes of average per capita income losses across the German federal states.³³ The southern states Bavaria (Bayern) and Baden-Württemberg not only face the most severe heat load, but also have the highest total GDPs and high per capita incomes. Consequently, the per capita income losses reach around 50 € per person and year in Bavaria and Baden-Württemberg, while the mean income reductions in Schleswig-Holstein and Mecklenburg Western Pomerania (Mecklenburg-Vorpommern) amount to ca. 5 € p. a. and person only. Although the heat load is rather low in the northern cities Hamburg and Bremen, their estimated per capita income losses are about the German average loss of ca. 24 € p. a., because the per capita incomes in these northern cities are relatively high.

The results have a similar magnitude as cost estimates of around 10 billion € for Europe for the heat summer 2003.³⁴ Nevertheless, so far the economic impacts of heat waves have not been evident in countries' growth rates.³⁵ According to our results, heat already has a high negative influence on the German economy today

³³ Population data for 2004 from Statistisches Bundesamt (2005), Bevölkerungsfortschreibung; graphic based on a map of Germany from Universität Trier (2007).

³⁴ Michael Heise, chief economist of the Allianz Group, cited in *Welt am Sonntag* (23.07.2006), this number includes other economic effects of heat besides health impacts.

³⁵ Michael Heise and Claudia Kemfert, cited in *Welt am Sonntag* (23.07.2006).

which will substantially sharpen in the future. The main caveat is that we do not take into account who is affected by heat to what extent and who is not, and that such a disaggregated view is hard to predict for the end of the 21st century. Maybe most people will work in air conditioned environments, and machines will do any physical work, so that heat will have little influence on labor output.

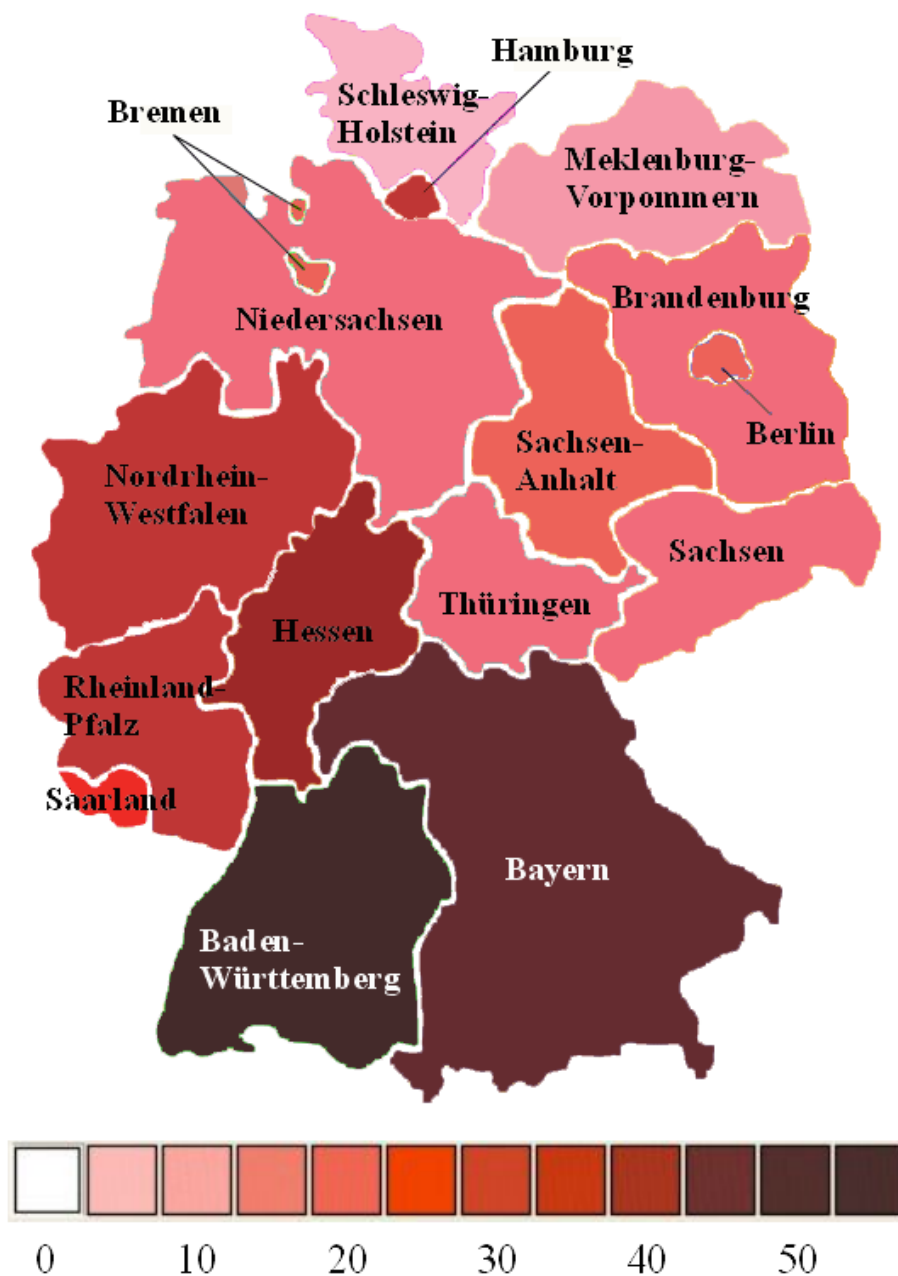


Figure 10: Heat related income losses p. a. and per capita, 2071-2100 across German federal states in 2004-€

5 Conclusion

Climate change will lead to a number of detrimental health impacts. In this study we focus on the effect of heat load on human well-being. The study is based on high resolution computations of climate data for the period 2071 to 2100 in Germany referring to IPCC scenarios. We calculate the change in the number of days with thermal stress. For this purpose, we apply perceived temperatures including humidity, wind and radiation, which better describe the heat subjectively felt by human beings than normal absolute temperatures. The results are then used to derive rough estimates of heat related health problems: the potential change in mortality, the potential costs of hospital emergency admissions, and the macroeconomic costs of a reduced work performance.

Heat induced mortality is examined as an indicator for the future health risk, since in contrast to particular heat related diseases sufficient statistical information on mortality is available. We expect an increase of heat related casualties by a factor of 3.7 at the end of the century compared with today. This increase is partly due to the rising share of elderly people who are known to suffer most from heat waves. The negative effects of heat waves, mainly in the south of Germany, dominate the reduction in the number of casualties because of milder winter periods by far. We do not take into account that a certain number of sick and elderly people might have died even without heat in the near future. Subtracting these casualties can possibly reduce the mortality numbers by up to 25 % (in the short-run).

The majority of negative health effects will not lead to deaths but may require medical assistance. Due to a lack of data we concentrate on heat induced hospital admissions. We find future hospitalization costs of 300 to 700 million € p. a. for Germany at today's prices, i. e. a 6-fold cost increase.

Finally, we look at the labor productivity of people under future heat load. A rough calculation yields a reduction of German GDP by 0.1 to 0.5 percent, i. e. a loss 4 times larger than at today's climate. These costs are significantly higher than the estimated hospitalization costs.

This study represents a first attempt to quantify the effects of climate change on health in Germany. The focus on direct temperature related effects was determined both by the size of the project and by the lack of conclusive results from the different disciplines concerned with the quantitative assessment of climate related health risks. The underlying empirical evidence of heat related hospital admissions and thermal

impacts on work performance shows a large variance. The results could become more robust if a larger research project brought together further interdisciplinary expert knowledge. A more detailed analysis of the working conditions in different sectors would enhance the quality of the estimates. Furthermore, more research is necessary to quantify the influence of climate change on health and productivity during the cold period of the year.

These estimates are all performed without considering physiological adaptation of the population and without including adaptation strategies in order to illustrate the size of the problem and because we did lack information on adaptation costs. Taking adaptation into account would change the numerical results and the appearance of the presented figures. One of the challenging tasks would be to assess different adaptation options and the related costs. While physiological and behavioural adaptation are hard to quantify in monetary terms, the cost of private and public adaptation measures (such as building design and urban planning) could be estimated. The simulations based on different IPCC scenarios also show that the health impacts can differ substantially depending on the emission path of greenhouse gases, thus indicating the benefits of mitigating climate change can be substantial. An analysis of the benefits of mitigation compared with the costs of adaptation would constitute a fruitful extension of the research that is presented here.

6 Acknowledgements

We thank the Umweltbundesamt for the REMO data, Deutscher Wetterdienst for providing the climate “Michel” model, Wolfram Mauser for contributing geographical population data and WWF Germany for initiating and supporting the project. Special thanks to Birger Tinz for computing and documenting the climate forecasts and for his competent help. We also thank Gerd Jendritzky and the anonymous referees for very useful comments.

7 References

- Basu, R. and Samet, J. M. (2002). Relationship between Ambient Temperature and Mortality: A Review of the Epidemiologic Evidence. *Epidemiologic Reviews* 24(2).
- Becker, P. (2005). Das Hitzewarnsystem des Deutschen Wetterdienstes: Notfallvorsorge. *Zeitschrift für Katastrophenmanagement und Humanitäre Hilfe*, 22-23.
- Bux, K. (2006). Klima am Arbeitsplatz – Stand arbeitswissenschaftlicher Erkenntnisse – Bedarfsanalyse für weitere Forschungen,. Bundesamt für Arbeitsschutz und Arbeitsmedizin, Dortmund, Forschung Projekt F 1987.
- Calado, R. M. D., da Silveira Botelho, J., Catarino, J. and Carreiro, M. (2005). Health Impacts of the 2003 Heat-Wave in France. In: W. Kirch, B. Menne and R. Bertollini, *Extreme Weather Events and Public Health Responses*. WHO Europe, EU Kommission, EEA, EUPHA, Springer Berlin Heidelberg New York, 89-97.
- Conti, S, Meli, P., Minelli, G., Solimini, R., Toccaceli, V., Vichi, M., Beltrano, M. C. and Perini, L. (2005). Epidemiologic Study of Mortality During Summer 2003 in Italian Regional Capitals: Results of a Rapid Survey. In: W. Kirch, B. Menne and R. Bertollini, *Extreme Weather Events and Public Health Responses*. WHO Europe, EU Kommission, EEA, EUPHA, Springer Berlin Heidelberg New York, 109-120.
- Donaldson, G., Kovats, R. S., Keatinge, W. R. and McMichael, A. J. (2002). Heat- and cold-related mortality and morbidity and climate change. *Health Effects of Climate Change in the UK*, Department of Health.
- EEA, European Environment Agency (2004). *Impacts of Europe's changing climate – An indicator-based assessment*. EEA Report No 2/2004.
- Fanger, P. O. (1972). *Thermal Comfort. Analysis and Applications in Environmental Analysis*. McGraw-Hill, New York.
- Frost, D. B. and Auliciems, A. (1993). Myocardial infarct death, the population at risk and temperature habituation. *Int. J. Biometeorol* 37, 46–51.
- Gagge, A. P., Fobelets, A. P. and Berglund, P. E. (1986). A standard predictive index of human response to the thermal environment. *ASHRAE Trans.* 92, 709-731.
- Guest, C., Willson, K., Woodward, K., Hennessy, K., Kalkstein, L. S., Skinner, C. and McMichael, A. J. (1999). Climate and Mortality in Australia: retrospective study, 1970-1990, and predicted in five major cities. *Climate Research* 13, 1-15.
- Hassi, J. (2005). Cold Extremes and Impacts on Health. In: W. Kirch, B. Menne, and R. Bertollini, *Extreme Weather Events and Public Health Responses*. WHO Europe, EU Kommission, EEA, EUPHA, Springer Berlin Heidelberg New York, 59-67.

Havenith, G. (2005). Temperature Regulation, Heat Balance and Climatic Stress. In: W. Kirch, B. Menne, and R. Bertollini, Extreme Weather Events and Public Health Responses. WHO Europe, EU Kommission, EEA, EUPHA, Springer Berlin Heidelberg New York, 69-80.

Hübler, M. and Klepper, G. (2007). Kosten des Klimawandels: Die Wirkung steigender Temperaturen auf Gesundheit und Leistungsfähigkeit. Eine Studie für WWF Deutschland.

IPCC (2001). Climate Change 2001: Synthesis report. Summary for policymakers. <http://www.ipcc.ch/pub/un/syreneng/spm.pdf>.

Jacob, D., Andrae, U., Elgered, G., Fortelius, C., Graham, P.L., Jackson, S.D., Karstens, U., Koepken, C., Lindau, R., Podzun, R., Rockel, B., Rubel, F., Sass, H. B., Smith, R. N. D., Van den Hurk, B. J. J. M. and Yang, X. (2001). A comprehensive model inter-comparison study investigating the water budget during the BALTEX-PIDCAP period. *Meteorology and Atmospheric Physics* 77, 19-43.

Jendritzky, G., Menz, G., Schirmer, H. and Schmidt-Kessen, W. (1990). Methodik der räumlichen Bewertung der thermischen Komponente im Bioklima des Menschen (Fortgeschriebenes Klima-Michel-Modell). *Beiträge d. Akad. f. Raumforschung und Landesplanung* Bd. 114, 7-69.

Jendritzky, G., Staiger, H., Bucher, K., Graetz, A. and Laschewski, G. (2000). The perceived temperature: the method of Deutscher Wetterdienst for the assessment of cold stress and heat load for the human body. Internet Workshop on Windchill, April 3-7, 2000, Meteorological Service of Canada, Environment Canada.

Johnson, H., Kovats, R. S., McGregor, G.; Stedman, J.; Gibbs, M.; Walton, H.; Cook, L. and Black, E. (2005). The impact of the 2003 heat wave on mortality and hospital admissions in England. *Health statistics Quarterly* 25, 6-11.

Kalkstein, L.S. and Davis, R. E. (1989). Weather and human mortality: an evaluation of demographic and interregional responses in the United States. *Annals of the Association of American Geographers* 79(1), 44-64.

Kampmann, B. (2000). Zur Physiologie der Arbeit in warmem Klima. Ergebnisse aus Laboruntersuchungen und aus Feldstudien im Steinkohlenbergbau. Habilitationsschrift, Bergische Universität – Gesamthochschule Wuppertal.

Koppe, C., Jendritzky, G. and Pfaff, G. (2003). Die Auswirkungen der Hitzewelle 2003 auf die Gesundheit. *DWD Klimastatusbericht 2003*, 152-162.

Koppe, C. and Jendritzky, G. (2004). Die Auswirkungen der Hitzewellen 2003 auf die Mortalität in Baden-Württemberg. *Gesundheitliche Auswirkungen der Hitzewelle im August 2003*, Sozialministerium Baden-Württemberg.

Koppe, C. and Jendritzky, G. (2005). Inclusion of short-term adaptation to thermal stresses in a heat load warning procedure. *Meteorologische Zeitschrift* 14(2), 271-278.

- Koppe, C. (2005). Gesundheitsrelevante Bewertung von thermischer Belastung unter Berücksichtigung der kurzfristigen Anpassung der Bevölkerung an die lokalen Witterungsverhältnisse. Bericht des Deutschen Wetterdienstes DWD 226, Offenbach.
- Kosatsky, T. (2005). The 2003 European Heat Waves. *Eurosurveillance* 10(7), 148-149.
- Kovats, R. S., Hajat, S. and Wilkinson, P. (2004). Contrasting patterns of mortality and hospital admissions during hot weather and heat waves in Greater London, UK. *Occup. Environ. Med.* 61, 893–898.
- Kovats, R. S. and Jendritzky, G. (2006). The Impact of Heat on Health. In: B. Menne, and K.L. Ebi, *Climate Change and Adaptation Strategies for Human Health*. WHO Europe, 63-97.
- Landscan (2001). Global Population Database, Oakridge, TN: Oak Ridge National Laboratory, www.ornl.gov/gist/.
- Laschewski, L. and Jendritzky, G. (2002). Effects of the thermal environment on human health: an investigation of 30 years of daily mortality data from SW Germany. *Climate Research* 21, 91-103.
- Max-Planck-Institut für Meteorologie (2006). Stellungnahme des Max-Planck-Instituts für Meteorologie zu den Artikeln von Volker Mrasek „Peinlicher Fehler im Klimamodell“.
- Spiegel-online (2006). Prognose verhaselt, *Financial Times*, Deutschland, 16. Oktober 2006, http://www.mpimet.mpg.de/fileadmin/grafik/presse/Stellungnahme_LAST_161006.pdf.
- McMichael, A. J., Whetton, P., Hennessy, R., Nicholls, N., Hales, S., Woodward, A. and Kjellstrom, T. (2002). *Human Health and Climate Change in Oceania: A Risk Assessment*. Commonwealth of Australia 2003.
- Parsons, K. (2003). *Human Thermal Environments – The effects of hot, moderate and cold environments on human health, comfort and performance*. Second Edition, Taylor & Francis.
- Schönwiese, C.-D., Staeger, T., Trömel, S. and Jonas, M. (2003). Statistisch-klimatologische Analyse des Hitzesommers 2003 in Deutschland. *DWD Klimastatusbericht 2003*, 123-132.
- Schwartz, J., Samet, J. M. and Patz, A. (2004). Hospital Admissions for Heart Disease – The effects of Temperature and Humidity. *Epidemiology* 15(6).
- Semenza, J. C., McCullough, J. E., Flanders, W. D., McGeehi, M. A. and Lumpkin, J. R. (1999). Excess Hospital Admissions During the July 1995 Heat Wave in Chicago. *American Journal of Preventive Medicine* 16(4), 269-277.

Sozialministerium Baden-Württemberg (2004). http://www.sozialministerium-bw.de/de/Meldungen/57521.html?_min=_sm&template=min_meldung_html&referer=80177.

Staiger, H., Bucher, K. and Jendritzky, G. (1997). Gefühlte Temperatur. Die physiologisch gerechte Bewertung von Wärmebelastung und Kältestress beim Aufenthalt im Freien in der Maßzahl Grad Celsius. *Annalen der Meteorologie* 33, 100-107.

Statistisches Bundesamt (2004a). Kostennachweis der Krankenhäuser, Fachserie 12, Reihe 6.3, 1.1.

Statistisches Bundesamt (2004b). Gesundheitswesen, Grunddaten der Krankenhäuser, Fachserie 12, Reihe 6.1.1, 2.2.2.

Statistisches Bundesamt (2005). Bevölkerungsfortschreibung, Fachserie 1, Reihe 1.3, 3.2.

Statistisches Bundesamt (2006a). Genesis online, Ergebnisse der 10. koordinierten Bevölkerungsvorausberechnung (Basis 31.12.2001), <https://www-genesis.destatis.de/genesis/online/logon>.

Statistisches Bundesamt (2006b). Statistisches Jahrbuch, Gesundheitswesen, 9.1.3.

Statistisches Bundesamt (2006c). Volkswirtschaftliche Gesamtrechnung, 24.4, 2004.

Universität Trier (2007). Deutschlandkarte der Hochschulen, <http://www.uni-trier.de/uni/unis/deutschlandkarte/karteC.htm>.

Uphoff, H. and Hauri, A. M. (2005). Auswirkungen einer prognostizierten Klimaänderung auf Belange des Gesundheitsschutzes in Hessen. Hessisches Landesamt für Umwelt und Geologie (HLUG).

VDI (1994). Umweltmeteorologie. Wechselwirkungen zwischen Atmosphäre und Oberflächen; Berechnung der kurz- und langwelligen Strahlung. VDI-Richtlinie 3789, Blatt 2.

VDI (1998). Umweltmeteorologie. Methoden zur human-biometeorologischen Bewertung von Klima und Lufthygiene für die Stadt- und Regionalplanung. Teil I. Klima, VDI-Richtlinie 3787, Blatt 2.

Weisskopf, M. G., Anderson, H. A., Foldy, S., Hanrahan, L. P., Blair, K., Török, T. J. and Rumm, P.D. (2002). Heat Wave Morbidity and Mortality, Milwaukee, Wis, 1999 vs 1995: An Improved Response?. *American Journal of Public Health* 92(5).

Welt am Sonntag (2006). Schon zehn Grad mehr Hitze senken die Arbeitsleistung um rund 30 Prozent. 23. Juli 2006. Jost, S, Wüpper, G., Struve, A. and Finkenzeller, K.

WHO Europe (2005). Health and Climate Change: the 'now and how' – A policy action guide. <http://www.euro.who.int/document/E87872.pdf>.

Wyon, D. (1986). The effects of indoor climate on productivity and performance – A review. *WS and Energy* 3, 59-65.

Zebisch, M., Grothmann, T., Schröter, D., Hasse, C., Fritsch, U., Cramer, W. (2005). Klimawandel in Deutschland – Vulnerabilität und Anpassungsstrategien klimasensitiver Systeme. Potsdam-Institut für Klimafolgenforschung im Auftrag des Umweltbundesamts, *Climate Change* 08/05, ISSN 1611-8855, 122-136.

Curriculum Vitae

Personal details

Name	Michael (Georg) Hübler
Address	14467 Potsdam, Germany
Phone	+49-331-288-2519
Email	michael-huebler(...)web.de
Borne	1977 in Fulda, Germany
Nationality	German

Working experience and internships

12/2009-	Potsdam Institute for Climate Impact Research, researcher in the area Sustainable Solutions
03/2006 – 11/2009	Kiel Institute for the World Economy, researcher in the area The Environment and Natural Resources
12/2003 – 03/2004	Vauxhall Engineering Centre Ltd Millbrook, Millbrook Proving Ground Ltd, IBC Vehicles Factory Ltd Luton (England)
02/1999 – 04/1999	K + S Kali GmbH Neuhof, electrical workshop
09/1998 – 10/1998	K + S Kali GmbH Neuhof, mechanical workshop
08/1994, 04/1996	Rensch-Haus GmbH Kalbach, design office

Studies

04/2006 – 04/2010	Christian Albrechts University Kiel, International Doctoral Program Quantitative Economics, doctoral thesis: Global Economic Integration, Technology Diffusion and Climate Change
10/1998 – 03/2006	Darmstadt University of Technology, Wirtschaftsingenieurwesen Elektrotechnik (economic sciences and electrical engineering), specialization in international economics and microelectronics, diploma thesis in economics: EU Eastern Enlargement – Allocation of Labour and Capital
09/1995 – 08/1999	Institute for Sacred Music Fulda, C-exam for part time musicians

School and civilian service

07/1997 – 07/1998	Diakonisches Werk Fulda, mobile social service
09/1988 – 06/1997	Freiherr-vom-Stein-Schule Fulda, grammar school

Affiliations

03/2010- Kiel Institute for the World Economy, research affiliate

Scientific activities

Selected presentations Workshop ‘Globalization and Environment’ Kiel 2009; European Association of Environmental and Resource Economists Annual Conference Amsterdam 2009, Jahrestagung des Vereins für Socialpolitik Graz 2008, ISOS lecture series Cluster of Excellence ‘The Future Ocean’ Kiel 2007, Earth System Science Partnership Open Science Conference Beijing 2006

Reviewing Climate Policy, Ecological Economics, Environmental and Resource Economics, International Trade and Economic Development

Session organizing Global Economic Symposium Plön castle 2008 and 2009, session on climate and energy

Publications and working papers

Hübler, M. (2010). Technology Funds: When Does Subsidizing FDI Save Carbon Emissions? Working Paper, 22 pp.

Kretschmer, B., Hübler, M., Nunnenkamp, P. (2010). Does Foreign Aid Reduce Energy and Carbon Intensities in Developing Countries? Kiel Working Paper, 1598, Kiel Institute for the World Economy, 28 pp.

Hübler, M., Keller, A. (2010). Energy Savings via FDI? Empirical Evidence from Developing Countries. *Environment and Development Economics*, Cambridge Journals, 15(1), 59-80.

Hübler, M. (2009). Can Carbon Based Import Tariffs Effectively Reduce Carbon Emissions? Kiel Working Paper, 1565, Kiel Institute for the World Economy, 26 pp.

Ignaciuk, A., Peterson, S., Hübler, M., Dellink, R., Lucas, P., Hilderink, H. (2009). An economy model for GISMO: DART-PBL technical documentation. Netherlands Environmental Assessment Agency (PBL), 24 pp.

Hübler, M. (2009). Avoiding the Trap: The Dynamic Interaction of North-South Capital Mobility and Technology Diffusion. Forthcoming in *The Journal of International Trade and Economic Development*, 30 pp.

Hübler, M. (2009). Energy Saving Technology Diffusion via FDI and Trade: A CGE Model of China. Kiel Working Paper, 1479, Kiel Institute for the World Economy, 40 pp.

Hübler, M., Lontzek, T. (2009). The Optimal Transfer of Capital and Embodied Technologies to Developing Countries. Kiel Working Paper, 1478, 30 pp.

Hübler, M. (2009). The Labor Market Effects of Outsourcing Parts and Components: Adverse Cournot Competition. Kiel Working Paper, 1564, Kiel Institute for the World Economy, 11 pp.

Hübler, M. (2008). The Labor Market Effects of Outsourcing Parts and Components: A Simple Model with Cournot Competition. *Aussenwirtschaft*, 63(2), 167-194.

Hübler, M., Klepper, G., Peterson, S. (2008). Costs of Climate Change - The Effects of Rising Temperatures on Health and Productivity in Germany. *Ecological Economics*, Elsevier, 68, 381-393.

Hübler, M. (2008). Prognose und Bewertung der zukünftigen Hitzebelastung in Bayern. In: J. G. Gostomzyk & M. Enke (eds.), Globaler Klimawandel und Gesundheit, *Schriftenreihe der Landeszentrale für Gesundheit in Bayern*, 19, 113-125.

Hübler, M. (2008). Sozio-ökonomische Bewertung von Gesundheitseffekten des Klimawandels in Deutschland. In: J. L. Lozán, H. Grassl, G. Jendritzky, L. Karbe & K. Reise (eds.), Warnsignal Klima: Gesundheitsrisiken - Gefahren für Menschen, Tiere und Pflanzen, *GEO Wissenschaftliche Auswertungen*, Hamburg, 299-306.

Hübler, M., Klepper, G. (2007). Kosten des Klimawandels: Die Wirkung steigender Temperaturen auf Gesundheit und Leistungsfähigkeit. WWF Germany, Hamburg, 65 pp.

Prizes

- | | |
|-------------------|--|
| 02/2003 – 06/2003 | Music competition „Jugend musiziert“ (organ, age group VII), 1. prize on national level |
| 05/1986 – 06/1992 | Genossenschaftsbank painting competition for the youth, prizes on regional and federal state level |

Ich erkläre hiermit an Eides Statt, dass ich meine Doktorarbeit „Global Economic Integration, Technology Diffusion and Climate Change“ selbständig und ohne fremde Hilfe angefertigt habe und dass ich alle von anderen Autoren wörtlich übernommenen Stellen, wie auch die sich an die Gedanken anderer Autoren eng anlehnenden Ausführungen meiner Arbeit, besonders gekennzeichnet und die Quellen nach den mir angegebenen Richtlinien zitiert habe.

Potsdam, 20. April 2010

Michael Hübler