DISS. ETH NO. 21805

Essays on Applied Econometric Analysis Related to Macroeconomics

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Sciences

presented by
FILIPPO LECHTHALER
Master of Science in Business and Economics, University of Basel
7.12.1981
citizen of Val Müstair (GR)

accepted on the recommendation of
Prof. Dr. Lucas Bretschger
Prof. Dr. Massimo Filippini

2014

Preface

During the last four years I had the opportunity to meet and interact with several people whose influence has been essential for this work. First of all I would like to thank my advisors Lucas Bretschger and Massimo Filippini for supervising this thesis. I am especially indebted to Lucas for giving me the possibility to develop and realize my research ideas and, at the same time, providing me with thoughtful critique and guidance.

I am thankful to my colleagues from team RESEC at ETH Zurich as well as at the Faculty of Business and Economics at University of Basel: I especially thank Julien Daubanes, Tobias Erhardt, Giulia Felber, Lisa Leinert, Max Meulemann, Markus Roller, Peter Schmidt and Aryestis Vlahakis for their valuable comments, discussions and cooperations.

I would also like to thank Kaspar Wyss and Joao Costa from the Swiss Tropical and Public Health Institute at University of Basel for bringing me on board Project Sino and for providing useful guidance on how to apply economics to health system analysis. Working on practical economic topics and political consulting during my time as a doctoral student considerably broadened my research horizon. Special thanks also go to the whole Project Sino staff in Dushanbe Tajikistan for the welcoming working environment.

I would like to express my sincerest thanks and appreciation to my friends and family for numerous wonderful moments, which served as a necessary inspiration when sitting behind my desk.

Pontresina, January 2014 Filippo Lechthaler

Contents

List of Figures

List of Tables

List of Abbreviations

AB	Arellano-Bond (estimator)
ACF	Autocorrelation function
AIC	Akaike information criterion
ARCH	Autoregressive conditional heteroskedasticity
BB	Blundell-Bond (estimator)
BE	Between estimator
BP	Breusch-Pagan (test)
CAPM	Capital asset pricing model
CLI	Composite leading indicator
CPI	Consumer price index
CO_2	Carbon dioxide
DGP	Data generating process
DSGE	Dynamic stochastic general equilibrium
EW	Equally weighted (portfolio)
EIA	Energy information administration
FE	Fixed effects (estimator)
GDP	Gross domestic product
GNI	Gross national income
GMM	Generalized method of moments
HML	High minus low
HQ	Hannan-Quinn (criterion)
ICAPM	Intertemporal capital asset pricing model
IIP	Index of industrial production

IS-LM	Investment-saving / Liquidity preference money supply
JB	Lomnicki-Jarque-Bera (test)
LSDV	Least square dummy variable (estimator)
LTCB	Long term credit bank of Japan
MEI	Main economic indicators
OECD	Organisation for economic co-operation and development
OLS	Ordinary least squares
OPEC	Organization of petroleum exporting countries
PACF	Partial autocorrelation function
RE	Random effects (estimator)
SC	Schwarz criterion
SMB	Small minus big
SVAR	Structural vector autoregression
SYS-GMM	System generalized method of moments
TOPIX	Tokio stock price index
VAR	Vector autoregression
VW	Value weighted (portfolio)
WML	Winner minus loser

Thesis Summary

The empirical evaluation of macroeconomic concepts and theories relies on a broad variety of econometric tools. Different methodological approaches have been used to overcome difficulties in identifying causal effects in observed macroeconomic data. Typical problems for parameter estimation are imposed by a high incidence of simultaneity bias as well as imprecise data measurement. Accordingly, the field of applied econometric analysis related to macroeconomic research has frequently been subject to methodological discussions questioning the validity of existing work and, thus, highlighting the lack of credible empirical evaluation of theoretical findings.

The present thesis contributes to this debate in two regards: first, we apply econometric methodology for exploring different macroeconomic hypotheses. Various identification strategies and methodological problems are discussed in order to properly verify theoretical findings from macroeconomic research. Second, we apply simulation analysis in order to evaluate different estimators leading future researchers towards more consistent results.

The methodological discussion in this thesis is built around three broad topics which are all related to empirical macroeconomics: identification of economic structure in the analysis of time series data, parameter identification in the context of growth empirics, and structural stability. Based on a historical background, chapter 1 describes the emergence and development of these three general topics.

Chapter 2 evaluates the determinants of the crude oil price after the year 2003. Specifically the price peak in the year 2008 has triggered an intense

debate about its ultimate causes. The economic literature proposes different explanations reflecting fundamental market forces (current supply and demand) and forward-looking market activities. We use a structural VAR model to disentangle the corresponding effects on the price surge after 2003. We find that market activities based on expectations regarding future market conditions have played a dominant role for the price development.

Chapter 3 and 4 both deal with economic growth. Chapter 3 analyses energy as a determinant of long-run economic growth in a cross-country analysis. Different estimation strategies related to panel data analysis are explored in order to identify the corresponding causal effect. We allow energy use to influence economic growth through an input substitution effect, that is through capital accumulation. Results suggest that the energy input is especially important as a growth determinant for emerging countries. Chapter 4 looks at the role of economic growth for the Japanese stock market. More specifically, using regression analyses and different statistical tests related to structural stability, we find that the presence of different growth regimes - the high growth period until the 1990s and the stagnation period in the following years - have a crucial impact on the stock pricing process.

Chapter 5 provides a simulation analysis evaluating bias properties of estimators in the context of growth regressions. We apply bias-correction terms for the lagged dependent variable and compare bias performances with more conventional estimators. Results suggest that although the correction term improves estimation properties for the parameter related to the lagged dependent variable, it does not outperform the other estimators in term of overall bias.

Kurzfassung

Die empirische Auswertung makroökonomischer Konzepte und Theorien basiert auf einer grossen Vielfalt von ökonometrischen Methoden. Verschiedene Ansätze wurden zur Identifizierung kausaler Effekte aus makroökonomischen Daten eingesetzt. Typische Probleme der Parameter Schätzungen sind beispielsweise Verzerrungen aufgrund des Simultanitätsproblems sowie unpräzise Datenmessung. Das Themengebiet der öknometrischen Analyse in makroökonomischen Anwendungen ist dadurch vermehrt Gegenstand methodischer Debatten, wobei die Glaubwürdigkeit bestehender wissenschaflticher Beiträge in Frage gestellt und somit mangelhafte empirische Evaluation bestehender Theorien verzeichnet wird.

Die vorliegende Arbeit trägt auf zwei Weisen zu dieser Diskussion bei: einerseits verwenden wir ökonometrische Methoden um makroökonomische Hypothesen auszuwerten. Wir diskutieren verschiedene Identifikationsstrategien und methodische Probleme um entsprechende theoretische Erkenntnisse aus der ökonomischen Literatur empirisch zu verifizieren. Andereseits verwenden wir Simulationsanalysen um Schätzungsmethoden zu evaluieren. Die Resultate sollen künftigen Forschern zu genaueren Resultaten verhelfen.

Der methodische Inhalt dieser Arbeit basiert auf drei generellen Themen: die Identifikation von ökonomischer Struktur in der Analyse von Zeitreihen, die Identifikation von Parametern in Wachstumsregressionen und strukturelle Stabilität. Diese drei Theme werden im ersten Kapitel anhand eines historisch-methodolgischen Hintergrunds genauer vorgestellt.

In Kapitel 2 evaluieren wir die Determinanten des Erdölpreises nach dem Jahr

2003. Vor allem das Preishoch im Jahr 2008 hat eine intensive Debatte zu möglichen Einflussfaktoren ausgelöst. In der öknomischen Literatur werden dafür verschieden Erklärungen geliefert. Diese widerspiegeln grundsätzlich zwei Erklärungsansätze: fundamentale Marktfaktoren (gegenwärtiges Angebot und gegenwärtige Nachfrage) sowie vorausschauende ("spekulative") Markttätigkeiten. Wir verwenden ein strukturelles VAR Modell um die entsprechenden Effekte aufzuschlüsseln. Wir schlussfolgern, dass der Preisanstieg grösstenteils durch vorausschauende Markttätigkeiten verursacht wurde.
Kapitel 3 und 4 sind beides empirische Anwendungen zum Thema Wirtschaftswachstum. In Kapitel 3 untersuchen wir anhand eines Ländervergleichs was die Rolle von Energie für Wirtschaftswachstum ist. Wir benutzen verschiedene Schätzungsverfahren aus der Paneldaten-Literatur um den kausalen Effekt von Energie auf Wachstum zu identifizieren. Die Resultate zeigen, dass Energie als Produktionsfaktor lediglich für Schwellenländer eine entscheidende Rolle spielt und dass dort eine Reduktion des Energieverbrauchs auch Wachstumseinbussen zur Folge hat. Kapitel 4 betrachtet die Rolle von Wirtschaftswachstum für den japanischen Aktienmarkt. Wir verwenden Regressionsanalysen und Testverfahren zu struktureller Stabilität um aufzuzeigen, dass der Einfluss von zwei markanten Wachstums-Regimen in Japan (hohes Wachstum bis in die 1990er Jahre und danach Stagnation) einen erheblichen Einfluss auf die Preisbestimmungsprozesse im Aktienmarkt hat.
Kapitel 5 umfasst eine Simulationsanalyse, mit welcher wir verschiedene Schätzungsverfahren für Wachstumsregressionen auswerten. Wir verwenden ein Verzerrungs-Korrektur-Verfahren für die verzögerte abhängige Variable und vergleichen dessen Eigenschaften mit üblichen Schätzverfahren. Die Resultate zeigen, dass das Korrekturverfahren die Schätzgenauigkeit für die verzögerte abhängige Variable verbessert. Betrachtet man aber die durchschnittliche Verzerrung über alle Modell-Parameter, schliessen die üblichen Methoden besser ab.

Chapter 1

Introduction

This thesis consists of four self-contained essays that deal with different aspects of applied econometric analysis in the field of macroeconomics. In these essays we discuss and work on typical methodological problems by applying statistical models to macroeconomic hypotheses and by evaluating estimators based on simulation analysis. The fields of application include natural resource economics and economic growth.

The first essay employs a structural vector autoregressive (SVAR) model to study the determinants of the crude oil price after the year 2003. The second and third essays discuss the role of economic growth in different contexts: in the one essay we look at the role of energy use in a cross-country comparison and in the other essay we analyze the role of economic growth for the performance of the Japanese stock market. The last essay uses simulation analysis in order to compare different estimators for dynamic linear panel data models in the context of growth regressions. Based on the historical background of empirical macroeconomics, this first chapter describes the emergence of the methodological topics faced in the applications of this thesis along with its corresponding contributions.

1.1 Empirical Macroeconomics as a Distinct Research Field

The role and the credibility of empirical work in the economic discipline has been critically discussed in several notable works during the last half-century.[1] Essentially, the major methodological concerns are related to the statistical identification of model parameters as a central element in making causal inference from observed data.

Like theoretical economics, the empirical counterpart has been developed along two main disciplines: macroeconomics and microeconomics. The latter is mainly concerned with the analysis of data describing the individual behavior of single persons, households and firms. The former includes the study of fluctuations of aggregate measures of economic activity and prices as well as the determinants of long-run economic growth. As noted by Angrist and Pischke (2010), the field of empirical microeconomics has experienced vital progresses with regard to identification strategies. Credibility has increased mainly through the advances in the quality of empirical research designs using random assignments. This progress has been slower in empirical macroeconomics. In fact, a number of empirical macroeconomists abandoned classical econometric work in favor of computational experiments based on the Dynamic Stochastic General Equilbrium (DSGE) framework as described by Kydland and Prescott (1996). In the following I give a short description of the main developments and the corresponding challenges in the field of empirical macroeconomics.

According to Mankiw (2006), the Great Depression of the 1930s represents the starting point of macroeconomic research, motivating economists of the time to explain fluctuations of economic aggregates in the US market. *The General Theory of Employment, Interest and Money* written by John Maynard Keynes and published in 1936 was at the heart of the professional discussion

[1] See e.g. Hendry (1980), Sims (1980), Leamer (1983) and Angrist and Pischke (2010).

of the time. In the following years, often referred to as the *Keyensian Revolution*, many young academics provided concrete models explaining short-run fluctuations and the role of aggregate demand based on Keynes' work. The IS-LM model suggested by Hicks (1937) and Modigliani (1944) is a prominent example. It summerizes Keynes' idea by means of a simple mathematical structure. Simultanously, related econometric studies appeared, proposing models which were mainly used for forecasts and policy analysis in a macroeconomic context. By the 1960s several macroeconometric models were in use, all related to the Keynesian idea and the IS-LM structure, including hundreds of variables and equations. According to Mankiw (2006), the most prominent examples are the *Wharton Model* (Evans and Klein, 1967) and the *Data Resource Inc.* associated with Otto Eckstein.

A crucial development for empirical macroeconomics comes along with Lucas' (1973, 1976) contributions on the role of expectations, suggesting to rebuild conventional macroeconomic theory and the corresponding empirical analysis of the time. In terms of empirical work, Lucas' criticism is related to the parameters of prevailing large-scale macroeconometric model, representing only a reduced form, i.e. not being policy invariant. His critique suggested the introduction of *microfoundations*, that is, a theory of economic decision that would be invariant to changes in policy. The implications for empirical macroeconomics relate to modeling and estimating so-called *deep parameters* that are at the root of individual behavior such as preferences, technology and resource constraints. Thomas J. Sargent is frequently referred to as the main empirical representative of this *rational-expectation-movement*. In his works he combined the development of empirical methodology and concrete applications. In line with the Lucas critique, he formulated empirical models with microeconomic foundations taking active expectation formation into account. Once these structural parameters were estimated, he could use the models for evaluating policy experiments. His papers had a crucial impact on the re-interpretation of macroeconomic findings at that time, such as the role of monetary policy and the Philips-curve trade-off (see Sargent, 1971, 1973, 1976).

An alternative empirical approach arising from Lucas' conceptions and tightening macroeconomics to microeconomic foundations is the dynamic version of Arrow-Debreu's *General Equilibrium Theory* based on Kydland and Prescott's (1982) *Real Business Cycle Theory.* To date, the resulting methodological framework referred to as DSGE is a standard instrument in the field of macroeconomics for calibrating theoretical models to observed data and run policy experiments.

Sims (1980) introduced a further dominant concept for today's formal econometric work in macroeconomics. In his seminal contribution he questioned the validity of assumptions in existing macroeconometric models concluding that the resulting interpretations, forecasts and policy implications were shaky. The empirical models of that time typically consist of a system of simultaneous linear equations, where identification was based on the division of variables into those that are endogenous and exogenous. In particular, according to Sims (1980), it is hard to find "true" exogenous variables because expectations about macroeconomic outcomes are based on all available variables. He proposed a modeling concept imposing only little structure on the data. A VAR framework can be used to describe the relation between a set of endogenous variables and their lags. In a second step, the model parameters can be derived exploring different ways of identification. This so-callad structural VAR model has become a standard tool of modern macroeconometric analysis.

The topics described so far all deal with empirical strategies for analyzing short run fluctuations of measures of economic activity and prices. A further essential field of study in macroeconomics is concerned with the phenomenon of long-run economic growth. Important theoretical contributions were provided by Solow (1956) suggesting an analytical framework built around diminishing returns to capital and by Romer (1990) extending Solow's approach by endogenizing technological development.[2] The empirical analysis of economic growth started to gain ground in the late 1980s and became an industry since the 1990s. Early works by Baumol (1986) and Barro (1991) were followed by studies providing methodolgical advancement (see Mankiw, Romer and Weil 1992 or Islam, 1996) as well as a large number of applications. According to Mankiw (2006), the emergence and popularity of growth empirics in the 1990s was due to four reasons: First, the series of influential papers by Paul Romer (1986, 1990) offered a new set of ideas and tools for analyzing the large gap between poor an rich countries. Second, the new availability of cross-country data allowed for systematic studies of growth determinants. Third, the long and historically unique expansion of the US economy in the 1990s motivated many growth-studies. Last, the ongoing tensions in the macroeconomic disciplin between new classical and new keynesian views regarding short-run economic phenomena dragged many applied macroeconomists into the field of long-run economic growth.

Although empirical work in macroeconomics experienced important methodological advancements, today the discipline still suffers from a lack of credibility. In general it appears to be difficult to identify parameters in situations where typically a great deal of simultaneity is present. Several authors such as Summers (1991), Mankiw (2006), Solow (2008) and Angrist and Pischke (2010) criticize the excessive distance between theoretical work and the empirical counterpart as well as the validity of predominant identification strategies. Thus, it is rewarding to further explore, evaluate and develop methodological tools aiming at extracting empirical results in order to evaluate existing theories and provide forecasts and policy conclusions.

This dissertation contributes to this issue in two regards: on the one hand, it provides evidence for existing macroeconomic concepts aiming at narrowing the gap between theoretical and empirical macroeconomics. On the other hand, it evaluates different estimators frequently used in the field of

[2]Further vital contributions in the field of theoretical economic growth are Rebello (1991), Lucas (1988) and Romer (1986).

growth empirics leading researchers towards less biased model parameters and thereby more consistent policy implications.

1.2 Three Selected Problems

The field of empirical macroeconomics includes a broad variety of economic subjects and methodological approaches. The general topics of interest include inflation, unemployment, economic growth, the business cycle as well as monetary and fiscal policy. In this thesis I address two of these thematic fields: economic growth and business cycle analysis related to the development of the crude oil price.

In the last 70 years, a large set of methodological tools have been proposed in order to analyze marcoeconomic topics. This thesis will use two approaches: the structural VAR framework and the growth regression framework. The elaboration of identification strategies revealing causal relationship between economic measures is at the heart of this thesis. In this regard, three broad methodological problems can be identified. They are discussed along with possible solutions in the following sub-sections.

1.2.1 Time Series Analysis and Economic Structure

The first methodological concern relates to the area of applied macroeconomics dealing with the study of business cycle fluctuations in aggregate measures of economic activity and corresponding prices over short periods, i.e. times series data based on daily, weekly or quarterly variation. As we know from the existing literature, the fluctuations of many macroeconomic variables appear to be interrelated in an economically meaningful way.

From the historical context we have noted that the methodological strategies identifying economic structures behind observed macroeconomic data have caused intensive debates in the last century. With regard to higher frequency data, the problem, in essence, boils down to the specification of a suitable statistical technique in order to pin down the direction of causal relationships in situations where they are not obvious a priori.

The previous section shows, while the traditional approach to structural macroeconometric modeling is based on a system of dynamic simultaneous equations, such models have been criticized due to their need of exogenous variables (i.e. instruments) for identification. SVAR models provide an alternative as they consist of endogenous variables only and, thus, do not require exogenous variables for identification. In order to recover the structural parameters, identification in a SVAR model generally relies on restrictions imposed on the contemporaneous interplay of the variables under consideration. Sims (1980) suggested a specific recursive system. Alternative identification strategies have been proposed in the following years (see for example Sims (1986) or Blanchard and Quah, 1989). Empirical results can be derived by modeling and analyzing the unobserved structural shocks using impulse-response functions and cumulative effects of these shocks on the variables of interest. The model separates unexpected movements in the macroeconomic variables that can be viewed as fundamental causes of macroeconomic fluctuations. The impulse-response functions, in turn, disclose the dynamic impact of these shocks on the development of all the macroeconomic variables under consideration. In line with the basic idea of studying business cycle fluctuations, the analytical emphasis lies on the dynamic development of an economic system focussing on short-run transitions representing structural economic interrelations. Although SVAR models are frequently used to draw inferences in empirical monetary economics (see e.g. Bernanke and Blinder, 1992), the approach is increasingly applied to study short-run fluctuations in other fields such as energy and natural resource economics (see e.g. Kilian, 2009).

1.2.2 Parameter Identification and Growth Regressions

As described in the last section, starting in the late 1980s, the empirical growth literature deals with determining the factors of the vast observed differences in per capita income across countries. Other then the macroeconomic field analyzing business cycle fluctuations, economists studying growth focus on the development of economies over the long run in the absence of transitory shocks. Thus, as we will see in the present section, the methodological framework is build around the analysis of cross-sectional variation - a field which is typically occupied by microeconomists.

Although the empirical literature on growth and convergence has experienced important advances in the last few years, the branch still suffers from major methodological drawbacks with regard to parameter identification. More specifically, as it is shown by Bond et al. (2001) or by Hauk and Wacziarg (2009), the challenges in statistically modeling economic growth are related to poor grounding of the estimated functional forms in economic theory. Moreover, prevailing results may be governed by unjustified claims of causality due to several reasons. The right-hand-side variables are typically endogenous and measured with error, reflecting the high degree of simultaneity in observed macroeconomic data. There are omitted variables which are correlated with explanatory variables but not directly observable, e.g. the initial level of a country's efficiency which is expected to be associated with income growth. Furthermore, the small number of available observations complicates convergence properties of the commonly applied estimators.

To date, the unique solid theoretical fundament of empirical growth modeling is the framework provided by Solow (1956). Accordingly, Mankiw et al. (1992) firstly derive the functional form for cross-sectional estimation based on a aggregate production function. Growth rates are regressed on the log of initial income and a set of steady-state determinants. In line with the concept of conditional convergence, countries experience convergence after controlling for specific characteristics. Islam (1995) reformulates the model structure suggested by Mankiw et al. (1992) into a dynamic panel data model, representing the standard approach to empirical growth analysis to this day. He provides an extended aggregate production function by including an unobservable country-specific effect. Usual panel data procedures allow to account for these country-specific effects, addressing the problem of omitted (time-invariant) variables, which are potentially correlated with explanatory variables. Based on the dynamic panel data structure, Caselli et al. (1996) focus on Generalized Method of Moments (GMM) estimators which exploit the dynamic nature of the growth regression and use lagged variables as instruments as suggested by Arellano and Bond (1991). Accordingly, these instruments potentially allow for addressing the variety of endogeneity biases described above. However, as it is shown in chapter 5, insufficient finite sample properties of GMM estimators represent an additional source of consistency-problems.

In the last decade better research design helped at statistically identifying causal effects of economic growth. Acemoglu et al. (2001) use differences in mortality rates of European settlers in different colonies as an instrument for political institutions in the successor countries in order to address the problem of reversed causality between good institutions and high income growth. Applying similar strategies, Rodrick and Wacziarg (2005) and Persson and Tabellin (2008) evaluate the relationship between democracy and economic growth. Using instruments for identification, Bretschger (2010) introduces energy as a production factor in the empirical analysis of long-run economic growth.

We see that the challenge of finding credible identification strategies for observational data also governs the econometric debate around the growth regression. Researchers are thus required to carefully evaluate methodological trade-offs with regard to their specific applications.

1.2.3 Structural Stability in Dynamic Econometric Models

Econometric analysis of marcoeconomic topics typically involves data analysis based on variation over time or a combination of cross-sectional and longitudinal variation. Structural stability is a major issue in time series analysis. Parameter estimations are based on the assumption of stationarity which implies constancy of parameter moments. In case this assumption does not apply, inferences may be biased and forecasts lose accuracy. In fact, Stock and Watson (1996) show that parameter instability prevails in a substantial number of models constructed by 76 representative US monthly time series. Econometric analysis provides a rich pool of methods to identify different aspects of structural instability.[3] In the application of the present thesis, I focus on the following viewpoint: according to Hansen's (2001) description, a structural break occurs, if at least one model parameter has changed at some date (the breaking date) in the sample period. In terms of formal statistical testing, this problem is related to two developments in the structural change literature: the test of structural break of unknown timing, and the estimation of the timing of structural change. The former is based on Chow (1960), who proposes splitting the sample into sub-periods, estimate the parameters for subperiods, and then test for equality of the parameters applying F statistics. The latter provides a framework to identify the exact date when the change occurred. The breaking date is treated as an unknown parameter and is estimated applying the least square principle.[4]

Given the evidence for structural instability in a large number of existing time series studies, empirical macroeconomic research requires careful treatment of the above mentioned issues.

[3] Hansen (2001) or Zeileis et al. (2005) provide a comprehensive overview on the methodological state.

[4] A detailed methodolgical description is provided in chapter 2 and 4.

1.3 Content of this Dissertation

This thesis consists of four essays which are related to the methodological discussion in the present chapter. A brief overview is provided in the following.

Moody Oil - What is Driving the Price of Crude Oil?

The study in chapter 2 tries to clarify the determinants of the crude oil price after the year 2003. The price of crude oil is set in the global market and is therefore simultaneously determined with other macroeconomic aggregates.[5] Thus, the analysis of the development of the crude oil price naturally requires a macroeconometric approach in order to model the underlying economic structure and dynamic properties of the global crude oil market correctly. As we are interested in the short-run interactions of the crude oli price and real economic activity, we apply a SVAR model to take account of the corresponding methodological problems.

We pay particular attention to the role of expectations by including a time series for news items into the SVAR model, approximating anticipative market activities. Results show that shocks to market expectations, along with the increasing demand from emerging countries, have played a crucial role for the price surge after the year 2003.

Econometric analysis related to structural stability suggests that the time series model under consideration exhibit a break around the year 2003 representing the beginning of the prominent oil price peak in 2008. We show that ignoring structural instability crucially affects the model's results.

Energy Use and Economic Growth

As it is shown by Lindenberger and Kummel (2002), in conventional economic theory, energy's share in total factor costs is assumed to be of minor importance compared to labor or capital. However, the recessions after the oil

[5] See Barsky and Killian (2002) or Hamilton (2003) for an overview on the topic.

crisis in the 1970s and 1980s raise the legitimate question of how this minor importance can translate into such large economic impacts. Correspondingly, in chapter 3, we investigate the role of energy use as a factor for long-run economic growth.

The identification of the causal effect of energy use on economic growth from observed data is complicated through the presence of reversed causality. Higher energy use is expected to increase a country's production and a state of high income, in turn, is expected to cause a greater need for energy. We use the empirical growth framework proposed by Islam (1995). Results are explored by applying a variety of estimation approaches typically used for panel data. The problem of reversed causality is addressed by using the GMM approach applying instruments constructed by lagged variables. Results suggest that the interrelation between energy use and economic growth critically depends on the country's stage of development. We show that especially middle income countries which are at a transitional stage of economic development rely on energy use in order to grow.

Stock Performance and Economic Growth - The Japanese Case

Structural stability is at the center of the study presented in chapter 4. By applying econometric analysis of structural change to the Japanese economy we are able to combine two well-known observations related to economic growth and stock performance. First, in the last 50 years Japan experienced two distinct growth regimes: starting from the 1960s, a period of high growth and catch-up to the economies of the time, and, starting from the 1990s, a period of long-lasting stagnation. Second, with regard to stock performance, previous studies on the Japanese financial market find that evidence for the momentum strategy is weak (Liu and Lee, 2001).

We find that the conventional risk factor models (CAPM, Fama-French three-factor model and Carhart four-factor model) applied to the Japanese stock market exhibit structural instability and identify the breaking date in the late 1990s representing the change from the high growth regime to the non-growth regime. Accordingly, in line with Liew and Vassalou (2000), we show that risk-factors in Japan are significantly associated with economic growth and thus contain information related to the macroeconomic environment. Applying the risk-factor models to corresponding sub-samples (high-growth period and stagnation period) we are able to recover evidence for the momentum strategy. Hence, we conclude that stock pricing properties are linked to underlying growth regime in a crucial way. The Japanese case illustrates the necessity to account for structural stability when dealing with time series models over a long time horizon.

Addressing Biases in Dynamic Linear Panel Models

The growth regression derived by Islam (1995) is a special case of a dynamic linear panel model. In order to deal with individual- (country-) specific effects, parameter identification is often based on the Least square dummy variable (LSDV) estimator.[6] It is well-known, however, that this estimator does not converge to its true value in dynamic model specifications with a finite number of observations over time (Nickell, 1981). In order to empirically capture the process of long-term economic growth, data are usually averaged over five years, critically restricting data availability over time and, thus, potentially augmenting the incidence of bias.

Chapter 5 evaluates this specific source of bias in the context of growth regressions using simulation analysis. More specifically, we apply a bias correction mechanism for the LSDV estimator originally proposed by Kiviet (1995) and compare its performance with the typically employed panel estimators for growth regression.

Results show that the bias correction mechanism helps to identify the parameter for the lagged dependent variable. However, in terms of overall bias,

[6]The LSDV estimator is generally considered one version of the fixed effects (FE) approach. In the remainder of this thesis we will use these terms interchangeably. In both cases we refer to the within estimation procedure and thereby do not explicitly estimate the individual-specific dummies.

finite sample properties of estimators exploiting a combination of between and within variation appear to be superior.

Throughout this thesis, all calculations and estimations are conducted in R (R Development Core Team, 2014).

1.4 Contribution of this Dissertation

As described above, the present thesis provides two sorts of insights: on the one hand it gives evidence for established macroeconomic concepts aiming at closing existing gaps between theoretical and empirical macroeconomics. On the other hand, it provides methodological evaluations with regard to applied econometric analysis in the field of macroeconomics aiming at extracting general conclusions for improved estimation properties. The most important findings are restated in the following:

Economic insights:

- The price increase of crude oil after the year 2003 is driven by a combination of aggregate demand (current need) and - to a greater extent - market activities based on expectations regarding future market conditions.

- The impact of energy use on economic growth does depend on a country's stage of development. Middle income countries also referred to as emerging countries do rely on energy use in order to grow. For high and low income countries a decrease in energy input does not appear to harm long-run economic growth.

- The unique macroeconomic development in Japan is reflected in the performance of the Japanese stock market. Accounting for different growth regimes improves the description of stock returns by the conventional risk factor models.

Methodological insights:

- In macroeconomic applications, structural changes can affect parameter values and the respective interpretations decisively. Accordingly, empirical results should generally be evaluated against the background of structural stability. Two examples illustrate this point:
  - The time series models of the crude oil market applied in chapter 2 are sensitive with regard to structural instability. Testing procedures suggest that a structural break occurred around the year 2003. If this breaking point is not accounted for, results appear to be fundamentally different.
  - Structural instability is present in the risk factor models for the Japanese stock market, with the structural break coinciding with the change from the high growth to the stagnation period. After accounting for structural stability, we find a more accurate model performance especially with regard to the momentum strategy.
- Bias properties of the LSDV estimator applied to growth regressions can be improved by introducing additive correction terms. In fact, by applying the bias correction procedure, the convergence parameters is estimated with the smallest bias throughout several estimation methods.
- However, in terms of overall bias, panel data estimators including between variation perform better in the context of growth regression. Especially the Random Effect (RE) estimator reveals a superior performance.
- Simulation analysis for growth regressions also show that small sample properties of the GMM estimator proposed by Arellano and Bond (1991) are critically affected by the weak instrument problem. Accordingly, the estimation performance is improved considerably when including additional instruments as proposed by Blundell and Bond (1998).

Chapter 2

Moody Oil - What is Driving the Price of Crude Oil?*

The unparalleled surge of the crude oil price after 2003 has triggered a heated scientific and public debate about its ultimate causes. Unexpected demand growth particularly from emerging economies appears to be the most prominently supported reason among academics, suggesting that market participants did not anticipate future market conditions. We study the price dynamics after 2003 in the global crude oil market using a structural VAR model, paying particular attention to anticipative market activities. These are inferred from a time series of news items measuring the flow of publicly available information relevant for the crude oil market. We find that such forward-looking demand activities - instead of demand arising from real economic activity - have played an important role for the run-up in the price of crude oil after 2003. This indicates that market participants have anticipated a higher demand in the future, rather than having reacted to unexpected shocks from the current business cycle. We additionally find that emerging economies have not majorly contributed to the price surge.

* This chapter represents joint work together with Lisa Leinert.

2.1 Introduction

The price development in the crude oil market during the first decade of the new millennium has attracted attention due to mainly two reasons: first, the price rose from its low levels during the eighties and nineties over several years, taking on record heights in July 2008. This price development triggered an extensive debate about the "Third Oil Crisis", named after the two oil crises in the seventies and eighties. Second, and more importantly, in contrast to the first two oil crises, the reason for this recent price peak is not straightforward: Several potentially relevant developments took place simultaneously, complicating the identification of their effect on the price. The academic discourse usually moves around three explanations, reflecting the ultimate forces on the price of crude oil: First, it is claimed that the rising price reveals the finiteness of crude oil and the inability to further extent production capacities (supply-driven price increase). Second, it is hypothesized that the unexpectedly strong growth of emerging countries such as China and India has resulted in an unexpected increase of crude oil demand, leading to squeezes in the spot delivery of crude oil and a rising price (demand-driven price increase). Third, it is stated that the increasing number of speculators in the market of crude oil has considerably enforced the role of forward-looking demand activities and therewith altered the price dynamics (expectation-driven price increase).[1] Among the three explanations, the demand hypothesis has been averted most (Kilian (2009), Kilian and Murphy (2010), Kilian and Hicks (2012), Krugman (2008) and Hamilton (2008, 2009)). The supply hypothesis, as well as the expectation hypothesis have seen a less pronounced echo in the literature.

The major challenge in empirically assessing which of the three hypotheses

[1] While many factors potentially affect the price of crude oil at any point in time, we interpret demand and supply as primary forces on the spot price of crude oil. Focussing on these determinants implicitly assumes that other economic forces, e.g. interest rates and exchange rates, have a secondary effect on the spot price of crude oil.

provides a better explanation for the dynamics in the crude oil market after 2003 consists in isolating the different forces in its effect on the price. While price effects arising from supply are identifiable due to the ability of observing extracted quantities of crude oil, a differentiation between the remaining two potential causes of the price increase requires a careful decomposition of observed total crude oil demand into two "un-observable" parts: fundamental crude oil demand, i.e. demand for crude oil today that arises as a result of today's real economic needs for the commodity, and forward-looking demand which is triggered by the expectation of changes in the market of crude oil taking place in the future.[2] Both types of demand need to be approximated by suitable data.

As changes in the fundamental demand for crude oil mainly arise due to up- and downturns of the business cycle, it has been common procedure to represent it by appropriate business cycle indicators. However, an approximation of forward-looking demand is far less straightforward: such activities are driven by expectations and are not directly observable. Approximating forward-looking demand in empirical models on the oil market is still an open issue.

In this paper, we contribute to the question of what has driven the price of crude oil after 2003 by considering a new means of representing forward-looking demand. We use a time series of all news items related to the crude oil

[2]As crude oil is storable, it is possible to buy or sell units of crude oil in the future or spot market in expectation of future market conditions. Thus, the price reflects current conditions as well as expectations of future market conditions. Forward-looking or expectation-driven demand has been denoted as "precautionary demand" in other contexts (see e.g. Kilian, 2009) but similarly addresses demand activities related to an expected shortfall of supply to demand. Note that such forward-looking demand activities incorporate anticipative demand activities related to future real economic needs as well as demand activities anticipating future price movements but unconnected to real needs. Thus, we are able to identify the contribution of forward-looking demand activities to the price development (macro-level) but are unable to point down which type of market participant has caused the price increase (micro-level).

market that have appeared on news tickers of one of the world's largest news suppliers. This time series reconstructs the continuous flow of information to the crude oil market. It enables us to isolate the effect of informational innovations on the price of crude oil, beyond information contained in current supply figures and business cycle indicators. Such informational innovations are interpreted as expectation shocks as they may lead to an update of market participants' view on the future balance of supply and demand and may thus result in corresponding adjustment activities. For example, the news about the breakdown of an oil platform reports about an instantaneous shock to supply, reflected in current supply figures. In addition, this piece of news may lead to an update of expectations regarding the future balance of supply and demand, possibly leading to precautionary market activities. Thus, accounting for supply as well as the information flow allows to explicitly distinguish between the effect of a shock to market fundamentals and the effect of a shock to market participants' expectations on the price of oil.

With this proxy for forward-looking demand for crude oil at hand, we undertake a structural decomposition of the crude oil price in a VAR model. The methodology follows Kilian (2009). Results of the SVAR model show that shocks to expectations have played a crucial role for the price development after 2003.

As these results stand in contrast to previous contributions on this topic, we provide an extended sensitivity analysis in which we discuss possible triggers of our results, such as structural breaks in the time series, variations in the proxy for fundamental demand and the role of emerging economies for the global crude oil demand. In particular, we find evidence that structural breaks have occurred in the global crude oil market in 2003 which is crucial for the estimation results. Furthermore, we find no empirical support that current real economic activity in major emerging economies has driven the price surge after 2003. Thus, we conclude that the crude oil price development reflects the anticipation of future market fundamentals. In other words, it has been the expectation of future market conditions rather than unexpected shocks to current market conditions that explain the price movement.

Contributions investigating the trigger of the price development after 2003 can be roughly distinguished into two groups, depending on the methodology used to develop the results. A first class of studies relies on the identification of historic price patterns and interrelations with trader positions in order to investigate changes in the price formation (see e.g. Master, 2009, Tang and Xiong, 2011, Buyuksahin and Robe, 2011, Hamilton and Wu, 2011). A second class of contributions analyzes the price development in its macro-economic environment through the estimation of SVAR models (see e.g. Kilian, 2009, Kilian and Murphy, 2010, Juvenal and Petrella, 2011, Lombardi and van Robays, 2011, Morana, 2013) which has been a primary tool for econometricians to disentangle cause and effect of macroeconomic variables. Kilian (2009) is considered a seminal paper in this class of studies, proposing a framework for structural VAR analyses in the crude oil market. The aim of the paper is the identification of the relative contribution of supply, fundamental demand as well as precautionary demand to the historic oil price development. With respect to the events after 2003, he attributes growth in crude oil demand, caused by the strong and unexpected growth of major emerging economies, to be the main reason for the price increase. Accordingly, "markets were repeatedly surprised by the strength of growth" (Kilian and Hicks, 2012) with no role being played by supply or precautionary demand.

Given the persistent strong economic growth of emerging economies which lasted for almost a decade, it seems difficult to comprehend that market participants have not been able to adjust their expectations and market activities respectively. The choice of a proxy for expectation-based market activities may cause a crucial difference in the result whether the price increase reflects unexpected strong growth from emerging economies or the expectation of strong growth in the future.

Kilian and Murphy (2011), Morana (2013), Juvenal and Petrella (2011) and Lombardi and van Robays (2011) consider shocks to OECD crude oil inventories as a mean of capturing precautionary demand.[3] This approach, however, requires that OECD inventories data are correct, provided in a timely fashion and that they resemble activities of all market participants, including investment banks and growing economies such as China and India. However, especially in recent years, inventory data have become unreliable (Alquist and Kilian, 2007) and thus may constitute an imprecise proxy of the underlying economic variable (see also Singleton, 2011). The seminal paper Kilian (2009) did not model forward-looking demand activities explicitly but relegated all expectation-driven demand activities into the residuum.

Our information-based proxy relies on modeling the main trigger of expectation-based market activities, instead of using output-oriented proxies such as inventory data. Thereby we claim to be able to identify more directly the degree to which market participants have indeed adjusted their activities to the expectation of future market conditions, such as higher economic growth in emerging economies.

While Kilian and Hicks (2012) also use a particular type of news shock, revisions of GDP forecasts, in a structural VAR model, it is used to represent shocks to business-cycle related demand. This approach likely blurs the distinction between fundamental and expectation-based demand activities: Business-cycle forecasts provide the basis to form beliefs regarding future market conditions, equally affecting purchasing decisions of producers (real economic needs) as well as of purely profit-oriented speculators. Without additionally accounting for real economic needs, a precise distinction between both effects is not possible.

The chapter is organized as follows: Section 2.2 proposes the empirical model, the data and the results. Section 2.3 provides an extended discussion and Section 2.4 concludes.

[3]The rational is that "any expectation of a shortfall of future oil supply relative to future oil demand not already captured by flow demand and flow supply shocks necessarily causes an increase in the demand for above-ground oil inventories and hence the price of crude oil" (pg. 2).

2.2 The Empirical Model

In the following section, we propose a four-dimensional SVAR model for the time period of 2003-2010. The model incorporates an explicit differentiation between fundamental and forward-looking demand.

2.2.1 Model Description and Identification

The price of crude oil is set in a global market and is therefore simultaneously determined with other macroeconomic aggregates which complicates the identification process of the model's parameters. SVAR models provide a suitable approach in this context as they consist of endogenous variables only and, thus, do not require exogenous variables for identification. In return, the identifying strategy relies on restrictions imposed on the interplay of the variables under consideration. These restrictions typically cannot be tested and should therefore rely on a sound theoretical fundament. The empirical results are derived by modeling and analyzing unobserved structural shocks using impulse-response functions and cumulative effects of these shocks on the variables of interest.

Starting point for the estimation of an SVAR model is the estimation of its reduced form, i.e. a conventional VAR model, using OLS estimation methodology. The VAR model is based on monthly data for

$$\mathbf{y_t} = (prod_t, econact_t, sentiment_t, price_t)'$$

where $prod_t$ is the percentage change in global crude oil production, $econact_t$ refers to the economic activity index, $sentiment_t$ denotes the time series of news sentiment reflecting expectation based market activities and $price_t$ is the real price of crude oil. The number of lags, p, is chosen to be nine.[4] The VAR representation is

$$\mathbf{y_t} = \sum_{i=1}^{9} \mathbf{A_i y_{t-i}} + \mathbf{e_t}. \tag{2.1}$$

[4] See Section 2.2.6 for a justification of the choice of the number of lags.

The underlying SVAR models the contemporaneous effects between the variables $\mathbf{y_t}$

$$\mathbf{A_0 y_t} = \sum_{i=1}^{9} \mathbf{A_i^* y_{t-i}} + \varepsilon_\mathbf{t} \tag{2.2}$$

with $\mathbf{A_i} = \mathbf{A_0^{-1} A_i^*}$ and $\mathbf{e_t} = \mathbf{A_0^{-1}} \varepsilon_\mathbf{t}$.

The structural parameters cannot be identified without imposing restrictions on the model. While there are in general several techniques of how to impose such restrictions, we apply a parametric approach which is based on a recursive system.[5] We reduce the number of free parameters by imposing a triangular structure on the matrix $\mathbf{A_0}$. We impose the following restrictions:[6]

$$\mathbf{e_t} = \begin{pmatrix} e_t^{prod} \\ e_t^{econact} \\ e_t^{sentiment} \\ e_t^{price} \end{pmatrix} = \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{pmatrix} \varepsilon_t^{\text{flow supply shock}} \\ \varepsilon_t^{\text{flow demand shock}} \\ \varepsilon_t^{\text{news shock}} \\ \varepsilon_t^{\text{residual shock}} \end{pmatrix}. \tag{2.3}$$

In contrast to the reduced form disturbances $\mathbf{e_t}$ from the VAR model which are only linear combinations of the unidentified structural innovations $\varepsilon_\mathbf{t}$, residual shocks from the structural model can now be interpreted in a meaningful economic way. Flow demand and flow supply shocks represent unexpected changes in fundamental market forces whereas the news shock depicts changes in the forward-looking demand component. The SVAR parameters are determined using Maximum Likelihood methodology.

[5]Recursivity typically requires two types of assumptions: First, the structural shocks are assumed to be uncorrelated, i.e. the variance-covariance matrix Σ_ε is diagonal. The underlying economic interpretation is that the structural shocks do not have a common cause. Second, restrictions on the contemporaneous relationships of variables are imposed. Further methods for recovering structural parameters are long-run restrictions or sign restrictions. For more details on identifying restrictions see Fry and Pagan (2009).

[6]With the model being four-dimensional ($K = 4$), we set $\frac{K(K-1)}{2} = 6$ elements of matrix $\mathbf{A_0}$ equal to zero. The restrictions described in Equation (2.3) follow the justifications given in Kilian (2009).

2.2.2 News Sentiment as Estimate for Forward-Looking Demand

While current needs of crude oil contribute to total crude oil demand and thus to the price formation, the ability to store crude oil allows agents to act today to tomorrow's expected changes in the market of crude oil.[7] Thus, the spot price of crude oil contains views held in the market place regarding the future conditions of supply and demand in addition to current supply and demand conditions. Such expectation-based demand activities need to be explicitly modeled to correctly represent the relative contribution of each force to the price development.

A direct way of capturing expectations held in the market place consists of going to the roots of the expectation formation process. What affects the formation of expectations in the market? According to economic theory, the process is based on information that market participants receive over the course of time. Thus, a time series that captures in a continuous way all pieces of information that are relevant for the crude oil market is indicative for the expectations of market participants regarding the future development of supply and demand.[8]

[7] Expectations regarding the future development of supply and demand impact the price of crude oil through two channels: On the one hand, the price for a future delivery of crude oil can be agreed upon today on futures markets. On the other hand, crude oil is storable so that market participants can buy units today in anticipation of future market conditions. Thus, if an individual, for example, holds the expectation of a rising crude oil demand in the nearer future, she may take precautionary steps to avoid having to pay a high price in the future. She can either decide to buy a futures contract today (if the current futures price is still less than what she expects the spot price to be in the future) or buy crude oil today and store it. In both cases, her expectation of the future conditions of demand and supply will have an impact on the price of crude oil today, either via the futures market (and consequently, via the no-arbitrage condition also on the spot price) or via the spot market, directly.

[8] News have been used to model the formation of expectations in other contexts as well, see e.g. Lamla and Sarferaz (2012).

Two examples illustrate how news help to determine the forward-looking component in the price of crude oil. First, consider a news item reporting about an explosion of an important oil platform. This piece of news delivers information about a supply shock that will be visible in this month's supply statistics. In addition to this instantaneous supply shock, the piece of information also leads to an adjustment of expectations regarding the balance of supply and demand in the near future. It might thus also trigger adjustment activities of market participants today. In contrast, the news according to which OPEC announces to decrease supply in the near future affects the price of crude oil through only forward-looking market activities: Rather than waiting until the negative supply shock actually happens, market participants will react today to this piece of news. The price will have adjusted by the time the quantities actually change.

The Thomson Reuters News Analytics Database allows a re-construction of the continuous flow of information to the market. It contains all news items that have run over tickers in trading rooms. Time stamps characterizing the exact time of appearance of the news item as well as topic codes describing topics mentioned in the text allow for a selection of relevant news articles for the crude oil market and a construction of a continuous time series of news items. Due to the broad coverage the database is representative of the timely, public information available at least to professional investors, i.e. public news. The language used to describe the content, i.e. news sentiment, helps in quantifying the otherwise not quantifiable information of the news article.

Quantifying the content of a news article based on its language is a relatively new approach and has become possible through the advent of automated linguistic programs. The idea behind the program is that the overall tone of the language provides an indication of the expected movement of the underlying economic variable. For example, news articles reporting about an increase (decrease) of the economic variable referred to in the text naturally use more positive (negative) words. Thus, an article reporting about an increase (decrease) in supply or demand of crude oil can be expected to have been ascribed a positive (negative) sentiment. Articles reporting about an increase in supply or demand include news about an increase in OPEC supply, the finding of additional oil fields or an increase in world economic growth. In contrast, news articles reporting about a reduction in supply or a decrease in demand include articles on a war in resource rich countries, a reduction in the supply from OPEC countries, riots or strikes on oil platforms or upcoming economic recessions.

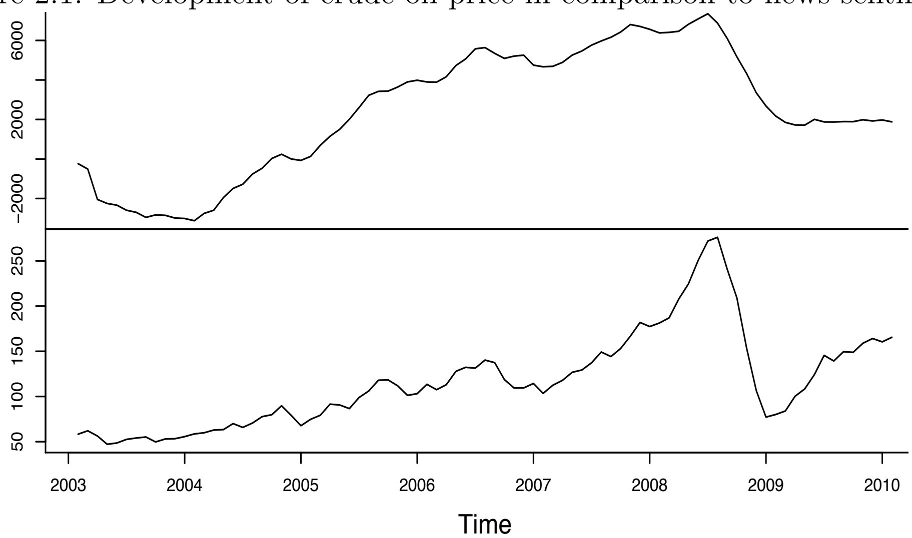
The sentiment attached to each news item is based on the tone of the language in each individual news article: On the basis of large dictionaries, the program counts the number of positive, negative and neutral words in each article and attaches a "1" ("-1" / "0") if the number of positive (negative/neutral) words outweighs the negative or neutral ones. Additional information on the likelihood of whether the sentiment variable correctly represents the tonus in the news article is given in form of probabilities ($prob_{pos}$ and $prob_{neg}$). The time series of daily sentiment is computed in the first step as

$$sents = \sum(1) \times prob_{pos} + \sum(-1) \times prob_{neg}. \tag{2.4}$$

A time series of monthly sentiment is given as the sum of daily news sentiments. Figure 2.1 shows the development of the crude oil price and the sentiment over time. Since 2003 the price of crude oil and the news sentiment have shown a high degree of co-movement: both, the time series of sentiment and the time series of the crude oil price, are increasing until the outbreak of the financial crisis and abruptly decreasing at the beginning of 2009. The years afterwards are characterized by a raising sentiment and price. The synchronous development of the two time series manifests itself in a high, positive correlation (0.815).

There are also several obstacles associated with using this time series as a proxy for market expectations. First, while the tone contains a signal regarding the expected change in supply or demand of crude oil, news items lack a reference to which economic variable they correspond to in particular.

Figure 2.1: Development of crude oil price in comparison to news sentiment



That is, we cannot observe a time series of news sentiment for supply and demand, separately. Still, we can derive some conclusions regarding the relative importance of supply- and demand-related news from descriptive statistics.[9] First, as we can observe whether news within a certain time period has been overly positive, negative or neutral and as we can observe the direction of the price, it is possible to ex-post infer the dominating type of news. As the correlation between the price of crude oil and news sentiment has been positive, it is clear that the time series cannot consist of a dominating number of references to supply. Furthermore, the news sentiment time series is highly correlated with the level of OECD production as measured by the index of industrial production provided by the OECD (0.765). Last, the application of a refined linguistic selection method further clarifies the importance of demand and supply related news. In order to approximate supply-related news sentiment (*supsent*) we filter news item based on the word OPEC representing exogenous changes in production which did not respond to current economic activity.[10] Demand-related news sentiment (*demsent*) is identified

[9]Note that estimated coefficients therefore will only reveal the average marginal effect of supply and demand-related expectations on the price of crude oil.

[10]Note that this is just one type of supply related news. A further refinement is work in progress.

by selecting news items related to economic indicators. The correlation as well as the multiple regression in Table 2.1 reveal that the supply-related news sentiment is not significantly linked to the general crude oil market news sentiment (*sent*). The demand related news sentiment, however, is significantly and positively linked to the crude oil sentiment which indicates that its content is governed by demand-related information. Based on the development of the sentiment and price time series, the question remains whether these news articles rather describe current conditions of the market or whether the news have resulted in the formation of certain expectations that were not accompanied by a corresponding shift in flow demand or flow supply. The SVAR decomposition where we account for fundamental supply and demand will allow for such a separation of effects.

The second issue regarding the use of news sentiment as proxy for market expectations relates to the question how many news items simply contain reports of the current development of the crude oil price, without containing further information on demand or supply. One could claim that the strong co-movement of the two time series and their high correlation is indicative of this hypothesis. The problem at the heart of this issue is the one of cause and effect between news sentiment and the price of crude oil. We shed some light on this causality by applying a bivariate Granger-Causality-Test. We find that, on a monthly basis, changes in news sentiment precede corresponding crude oil price dynamics.[11]

2.2.3 Motivation and Implications of Restrictions

The restrictions imposed on the contemporaneous relationships of the four variables described by equation (2.3) in Section 2.1 are explained in the following section.

[11] Note that this is work in progress for a separate paper. Preliminary results can be obtained from the authors upon request.

Table 2.1: News sentiment

Correlations			
	sent	*demsent*	*supsent*
sent	1		
demsent	0.194**	1	
supsent	0.085	0.383**	1

Regression: dependent variable *sent*		
	Coefficient	p-value
demsent	0.097	0.015
supsent	0.048	0.853

Notes: The estimations are based on weekly observations from June 2006 to Oct 2010.

Restrictions on Crude Oil Production

Shocks on crude oil production are amongst others caused by wars within crude oil producing countries, strikes on oil platforms, disruptions due to natural disasters as well as production regulation based on coordinated behaviour among OPEC members. This events typically do not react contemporaneously to demand shocks in the same month. Thus, adjustments of the production plan due to developments in the business cycle or the price of crude oil take place over a longer time horizon. As a consequence, we restrict production to be influenced in the *same* month by no other variable than a flow supply shock, itself ($a_{12} = a_{13} = a_{14} = 0$). The supply curve results to be vertical in the short run.

Restrictions on Fundamental Crude Oil Demand

Real economic activity (and thus the demand for crude oil associated with the business cycle) is affected in the same month by only a shock to the supply of crude oil or via a shock to the business cycle itself. Oil-market specific innovations such as shocks from oil-market specific news or the residual shocks will not affect global real economic activity immediately, but with a delay of at least 30 days ($a_{23} = a_{24} = 0$).

Restrictions on Sentiment

News sentiment (indicating expectation-driven demand) adjust to a flow supply shock, a flow demand shock as well as to a news shock in the same month. That is, we assume that market participants are capable of adjusting precautionary demand activities within 30 days after having learned about the outbreak of a war in resource producing countries or an upcoming economic crisis. Residual innovations not explained based on oil supply shocks, aggregate demand shocks or news shocks are excluded from affecting news sentiment within the same month ($a_{34} = 0$).[12]

The Price of Crude Oil

Last, the price of crude oil is the most reactive variable within the system as it responds instantaneously (i.e. within the same month) to flow supply, flow demand, news shocks and shocks that are not captured by any of the other three types of shocks (residual shocks).

2.2.4 Data

We use monthly global crude oil production taken from the Energy Information Administration (EIA) as measure of crude oil supply. The refiner acquisition cost of imported crude oil, deflated by the US CPI and expressed in logs, is taken as proxy for the real price of oil. We employ the index of industrial production as provided in the MEI database of the OECD as measure of business-cycle related crude oil demand. Last, we use the sentiment time series for the crude oil market as obtained from the Thomson Reuters

[12]Note that the bivariate Granger-causality test provides support for this restriction: changes in news sentiment precede corresponding crude oil price dynamics on a monthly basis.

News Analytics database, expressed in logs, as explicit measure for precautionary demand activities. The data run from February 2003 until February 2010. As the time series of the OECD production indicator as well as the one for crude oil production contain a unit root (see Table 2.3 in the Appendix), we transform the time series from levels into growth rates in order to achieve stationarity. While the crude oil price also exhibits a unit root, we restrain from any further transformation to preserve the information contained in the levels of the time series.

2.2.5 Results

Figure 2.2 represents the responses of the price to a unit shock from each of the four variables.[13] They allow four conclusions.

First, the responses of the crude oil price to a shock from flow supply and flow demand exhibit economically plausible patterns: A flow supply shock has a negative (not significant) impact on the price of crude oil after around nine months. A flow demand shock leads to a positive and significant increase in the price of crude oil, peaking after about eight months.

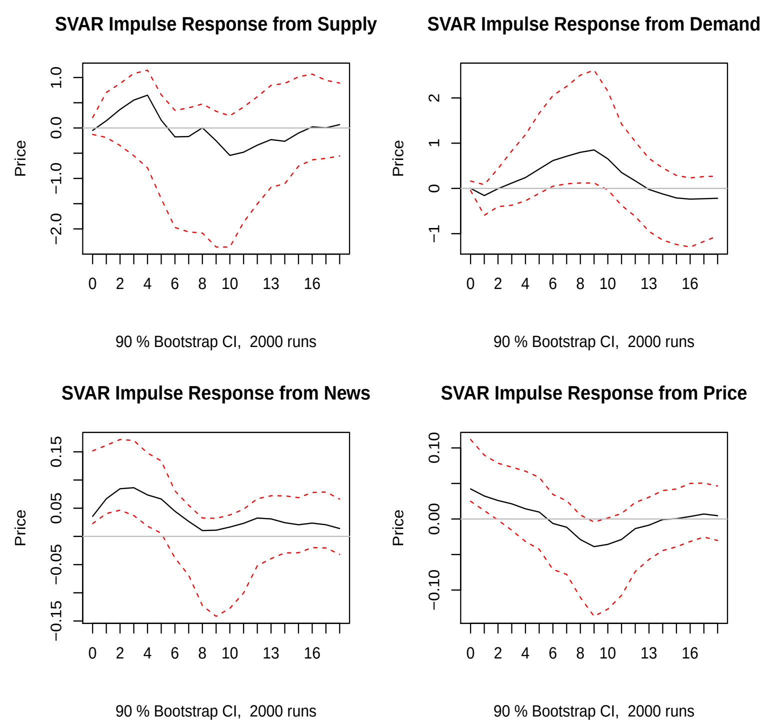
Second, a news shock has a highly significant and positive impact on the price of crude oil. The effect is significant from the impact period onwards and lasts for the following five months. A positive shock from news accordingly represents the expectation of higher demand in the future. This indicates that forward-looking demand activities have taken place, resulting in an increase in the price of crude oil. Note that a news shock does not have a reverting behavior of the price of crude oil. It remains positive over the course of the following 18 months.

Third, the results also indicate a reasonable difference in speed in the adjustment of the price of crude oil to flow demand and news shocks: while flow demand shocks arising from the business cycle need more than half a year to

[13] The impulse response functions for the response variables supply, demand and news are shown in the Appendix (Figure 2.11 to 2.13).

fully unfold their impact, news shocks have a rather short term impact on the price of crude oil with no significant influence after half a year.[14]
Last, residual shocks do also impact the price of crude oil significantly but show signs of self-reverting behavior. While a residual shock increases the price of crude oil significantly during the first two to three months, the shock turns negative over the following months.

Figure 2.2: Impulse response function for the Crude Oil Price

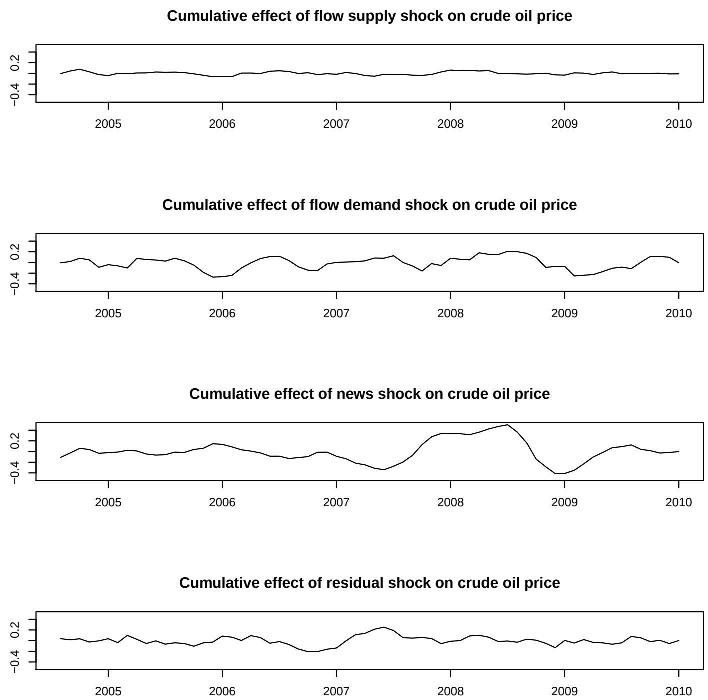


In accordance with the results from the impulse response functions, the historical decomposition of the price of crude oil in Figure 2.3 attributes most of the price development to news shocks. Especially around the year 2008, expectation-driven demand activities have influenced the price of crude oil in a notable way. While shocks from flow demand can explain some swings

[14]Adjustments to shocks from fundamental demand are clearly more sticky than adjustments to expectation shocks. While the latter does include costs from adjusting positions in the futures markets or adjusting inventories, the first incurs other costs, such as capacity adjustments.

in the price of crude oil, they did not contribute in a systematic way. Flow supply shocks did not help to explain the price development, at all.

Figure 2.3: Historical decomposition of crude oil price in four variables model



Last, cumulative effects from residual shocks are rather volatile but do not show a systematic pattern. This is in line with the overshooting pattern found in the impulse response function according to which the shocks did not have a persistent effect on the price of crude oil.

All in all, the results from the SVAR model do not support the hypothesis that unexpected shocks from real economic activity have caused the increase in the price of crude oil after 2003. Rather, they suggest that the price surge was mainly driven by news shocks, i.e. shocks to expectations regarding future market conditions.

2.2.6 Diagnostic testing

The results of an SVAR model do not only depend on the choice of the identifying assumptions but also on the specification of the underlying reduced-form model. Whereas the assumptions are imposed on the model on a priori grounds and cannot be tested directly, there are various statistical procedures for examining whether the reduced-form specification adequately represents the data generating process (DGP). In this section, following Breitung et al. (2004) or Pfaff (2008), we apply some well established diagnostic procedures. Figure 2.14 to 2.16 in the Appendix display the diagram of fit and the residual for every variable in the VAR model - flow supply, flow demand, news and crude oil price. Based on visual assessments, the plots of the residuals do not indicate any noticeable specification problems. In addition, the estimated autocorrelation function (ACF) as well as the partial autocorrelation function (PACF) for each single residual does not exhibit any significant deviation from zero at any lag.

In the following we apply multivariate tests to the model residuals. In a first step we test for the absence of autocorrelation. Two different procedures are considered: we perform a test based on an adjusted portmanteau statistic Q_h in order to check the null hypothesis of no autocorrelation against the alternative that at least one autocovariance is nonzero.[15] Secondly, as described in Godfrey (1978), we apply the Breusch-Godfrey LM (BP) statistic in order to test for hth order autocorrelation. As we can see from Table 2.2 both tests reject the null hypothesis of no autocorrelation in the model residuals. While a higher number of lags may reduce autocorrelation, the number of observations in our dataset imposes a severe trade-off in terms of asymptotic properties of the estimated parameters. Tests on optimal lag length (i.e. the Akaike information criterion (AIC), the Hannan-Quinn criterion (HQ) and the Schwarz criterion (SC)) indicate only little informational gain for lag three and beyond (see Table 2.2). In order to find a compromise between the optimal lag length, the autocorrelation patterns and the suggestion in previous papers of including long lag orders (e.g. Hamilton and Herrera (2004) and Kilian (2009)), we increase the corresponding number up to nine in order

[15] For a more detailed description of the following test statistics see Lütkepohl (2004).

to adequately represent the dynamics of the global crude oil market.
In a further step, we test for conditional heteroskedasticity in the error term by applying a multivariate extension of the univariate ARCH-LM test as described in Engle (1982). The corresponding p-value from Table 2.2 indicates that no ARCH effects are present.
Finally, we test for nonnormality in the error term. The test is based on the skewness and kurtosis properties and is constructed by generalizing the Lomnicki-Jarque-Bera (JB) test (Jarque, Bera; 1987). As we can see from Table 2.2 the null hypothesis of the residuals being normally distributed cannot be rejected. Based on the test results we conclude that the reduced-form model performs in a satisfactory manner, providing an adequate basis for the structural identification.

Table 2.2: Model checking: the VAR Specification

Diagnostic Tests							
Q_h	p-value	BG	p-value	ARCH	p-value	JB	p-value
164.029	0.001	147.866	$\leq$ 2.2e-03	512.845	0.336	2.250	0.972

Lag Length Selection					
AIC	lag length	HQ	lag length	SC	lag length
-28.416	2	-28.002	1	-27.634	1

2.3 Discussion

What has caused our results to differ so dramatically from those obtained in the reference literature? In the following section, we provide a discussion about possible factors causing the difference, e.g. the time period of estimation and the choice of the fundamental demand indicator. Last, we examine whether we can find empirical support for the hypothesis of demand from emerging economies triggering the price increase as it provides the backbone of the demand growth hypothesis.

2.3.1 Fundamental Changes and Structural Breaks

The first difference of our model in comparison to the estimations in the reference literature arises from the estimation horizon. We use data starting in 2003 due to the limited availability of the Thomson Reuters News Sentiment time series while many empirical assessments of the crude oil market use data over several decades. While longer time series are usually preferred as asymptotic properties of estimators improve with the number of observations, the likelihood of encountering structural breaks in the time series rises with the number of observations, as well. Ignoring the presence of such discontinuities when estimating a structural VAR model renders wrong parameter estimates and thus results. The finding of structural breaks occurring around 2003 would justify the concentration of our estimations on the shorter time horizon.

Various contributions have documented an altered functioning of the market for crude oil after 2003, indicating the likelihood of altered properties of the underlying time series (see e.g. Tang and Xiong (2011), Hamilton and Wu (2011)).

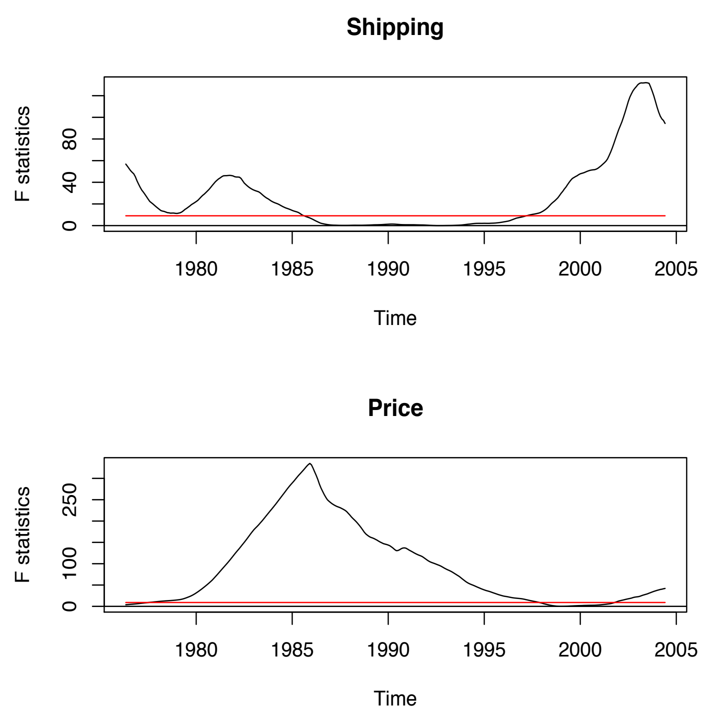
In order to find out whether central variables related to the market for crude oil have indeed experienced structural breaks within recent decades, we have applied a three-step test procedure to data most often used in the reference literature on oil price decompositions, i.e. Drewry's shipping index as proxy for business-cycle related demand, crude oil supply and the spot price of crude oil.

In the first step, we investigate for the three time series whether the mean differs significantly for sub-periods of the sample.[16] We compute an F-statistic in order to compare the unsegmented model against a possible break for each

[16]The data start in January 1973 and end in November 2007. For a detailed description of the data see Kilian (2009).

point in time. Following Andrews and Ploberger (1994), we reject the null hypothesis of structural stability if the supremum of these statistics is too large. We reject the null hypothesis of no structural change for the mean of shipping and the mean of the price at the 5% level. In order to see at which points in time the null hypothesis is rejected, we draw the process of the F-statistics for shipping and pricing (Figure 2.4), where the peaks roughly indicate the timing of possible structural shifts. The straight line illustrates the threshold for rejecting the null hypothesis. The process for shipping has three peaks: at the beginning in 1973, around 1982, and around 2004. The F-statistics for the price variable exhibit one peak around 1985 and one around 2003.

Figure 2.4: The process of the F-statistic



Given the evidence for structural instability for two time series, we assess the timing of the structural break following the procedure described in Bai and Perron (2003) in the next step.[17] We assume a three-segment partition with

[17] A more technical layout of testing procedure with regard to structural breaks is given in chapter 4.

two breaking points for both means based on the behavior of the F-statistics described above. The mean of the shipping variable contains breaking points in November 1981 and February 2003. The mean of the price variable occurs several months later, i.e. December 1982 and May 2003. Figure 2.5 illustrates the timing of the structural breaks and the mean for each sub-period for the time series of economic activity (”shipping”) and the real price of oil.

In a last step we find that the null hypothesis of no structural break in May 2003 is clearly rejected (test-value=932.910, p-value $\leq$ 2.2e-03) by applying a joint Chow test for structural breaks to the VAR model including the three variables under examination.[18]

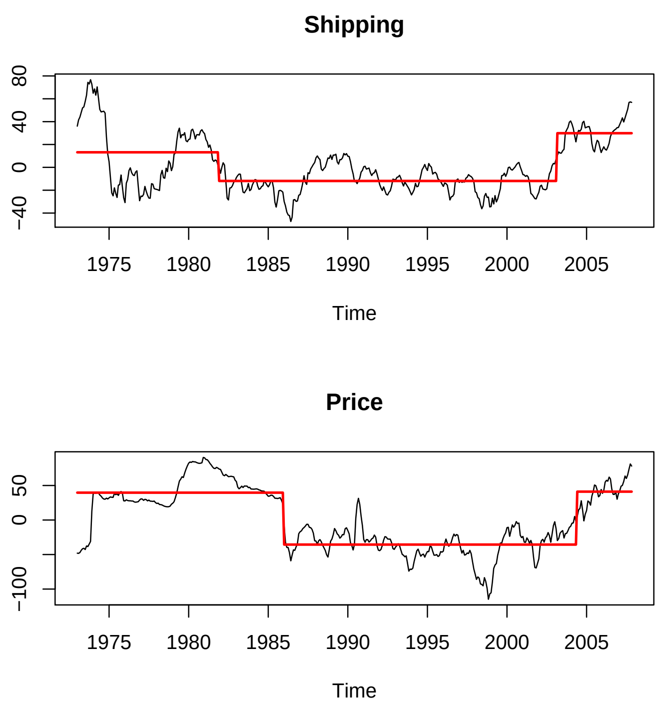
In summary, we find that the single time series of the spot crude oil price and of the economic activity indicator, as well as the SVAR model representing the global crude oil market, exhibit a structural break in 2003. This finding implies the need to focus on sub-periods for estimations that coincide with our chosen estimation period.[19]

The importance of acknowledging the presence of structural breaks in the estimations can be highlighted when re-estimating Kilian (2009) for the estimation period of 2003-2010. Focussing on this sub-period, results change dramatically: Not business-cycle related demand as in Kilian (2009), but forward-looking demand activities seem to have been the main driver of the price development of crude oil after 2003 (see Section 2.5.1 of the Appendix). Thus, the consideration of structural instability seem important in an empirical assessment of the crude oil market.

[18]For a detailed description see Lütkepohl (2004).

[19]Similar conclusions with respect to the occurrence of structural breaks are drawn in a variety of other contributions. Fan and Xu (2011) find three structural breaks which have occurred since the start of the new millennium: a ”relatively calm market” period (January 07, 2000, to March 12, 2004); the ”bubble accumulation” period (March 19, 2004, to June 06, 2008,); and the ”global economic crisis” period (June 13, 2008, to September 11, 2009). Further evidence of a change in the dynamics of the crude oil market is provided by Kaufmann (2011) who documents a structural break in the series on U.S. private crude oil inventories.

Figure 2.5: Breakingpoints and mean of sub-seriods



2.3.2 The Indicator for Business-Cycle Related Crude Oil Demand

A second source of variation of our model in contrast to the reference literature consists in the choice of the proxy for business-cycle related crude oil demand.[20] The reference literature has mainly used fundamental demand proxies based on shipping activities, i.e. the Baltic Dry Exchange Index or Drewry's shipping index, in order to infer crude oil demand associated with real economic activity. However, while one would expect an estimate

[20]It has become common practice to identify fundamental crude oil demand with the help of business cycle indicators. Such indicators are capable of indicating changes in the demand for crude oil that are purely based on an expansion or contraction of current world economic activity and thus demand for crude oil for today's use. Note that there are some crude-oil intensive activities that are not closely related to industrial production, e.g private traveling. However, such activities can be assumed to be highly correlated with the overall business cycle.

of business-cycle related crude oil demand to be correlated with either total crude oil demand or other indicators of the business cycle, we do not find a significant correlation with the commonly used estimators for the time period of 2003-2010: The correlation between the Baltic Dry Exchange and two alternative business cycle indicators, the index of industrial production (IIP) provided in the MEI database of the OECD and the Composite Leading Indicator (CLI) provided by the OECD, is -0.028 and 0.21, respectively, and not significant. In addition, the shipping indicator does not show any relation to figures on total crude oil demand, either (0.045, not significant).[21] Due to these obvious shortcomings, we have used the index of industrial production which is positively and significantly correlated with total crude oil demand (0.381).

The reference literature has refrained from using indices based on industrial production to proxy for crude oil demand associated with real economic activity due to several, presumed shortcomings (see Kilian, 2009). We argue, however, that they are not severe in the context of our estimations: First, the link between fundamental crude oil demand and industrial production figures has been found to be influenced by structural changes of economies and the development of new technologies. While this argument applies in particular to estimations conducted over several decades, it may be less problematic when investigating only a period of several years. Second, it has been argued that data on industrial production are only available for a fraction of countries in the world. For example, industrial production of major emerging economies are not yet contained in standardly available indices. Still, we find that countries for which data on industrial production is provided contribute on average by 77% to total world GDP between 2003 and 2010. China and India only contribute by 8% to world GDP.[22]

[21]The correlation-coefficients for all relevant variables are listed in Table 2.4 in the Appendix.

[22]However, looking at GDP increases only, emerging countries play a more prominent role: between 2003 and 2010 China and India contribute to global increase in GDP by an

Replacing the shipping index in the estimations of Kilian (2009) by the index of industrial production as provided by the OECD for the time period of 2003-2010, we find that the latter provides a better explanatory power for the price than the shipping indicator.[23] The results of the estimation are shown in Figure 2.9 and Figure 2.10 in Appendix 2.5.2.

The overall conclusion from the structural decomposition using the alternative proxy for fundamental demand remains the same: forward-looking demand shocks have mainly contributed to the price increase after 2003 (Figure 2.10). Thus the results obtained in Section 2.3.1 are robust to the choice of the demand indicator.

2.3.3 The "China-Effect" - The Role of Emerging Economies for the Development of the Crude Oil Price

As a last sensitivity analysis, we investigate the demand growth hypothesis in greater detail, i.e. the claim that demand from emerging economies such as China and India has driven the price increase in 2003.

We re-run the model in Section 2.2 but replace the industrial production indicator for OECD countries by two sorts of leading indicators: the first composes of only OECD countries, the second additionally includes major non-member economies (MNEs), including China, India, Russia, South Africa and Indonesia.

Note that due to the characteristics of a leading indicator these results will be only informative with respect to a *comparison* of cumulative effects of flow demand shocks. The results cannot be used as a comparison of the role of fundamental versus forward-looking demand for the price development as the leading indicator contains expectations regarding the development of the busi-

average of 30%, whereas the contribution from the OECD is 42%.

[23]The shocks from the industrial production indicator on the price appear to be partly significant in contrast to shocks from the shipping indicator. As in Section 2.3.1, the industrial production indicator is considered in growth rates rather than levels.

ness cycle. The leading indicators therefore capture part of the information contained in the news sentiment time series.

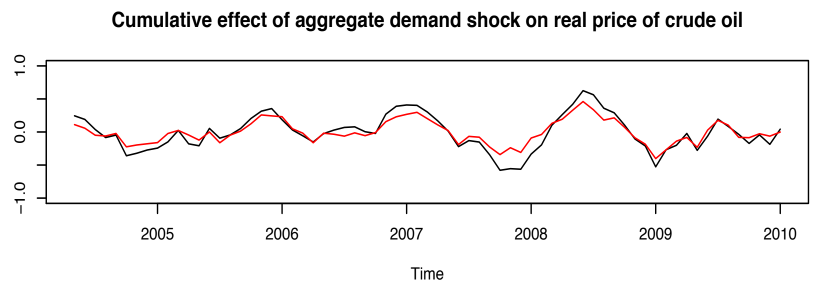


Figure 2.6: Comparison of cumulative effects for OECD and OECD plus major emerging economies in four variables model

Figure 2.6 shows the cumulative effects from fundamental demand on the price of crude oil, using the CLI for OECD countries and the CLI for OECD plus major non-member economies (MNE). The red line refers to the estimation based on the CLI of OECD countries (=benchmark case) whereas the black line refers to the CLI including major non-member economies. The graphs do not show a huge difference for the role played by fundamental demand in the run up of the price. The most notable difference arises in 2008, during the price peak, when cumulative flow demand shocks from OECD plus MNE countries on the price are slightly higher than those for only OECD countries. Still, considering the entire time period, emerging economies have not contributed to a large extent to the run up in the price of crude oil. Thus, we cannot find empirical support for the claim that the growth in emerging economies have majorly contributed to the price rise.

2.4 Conclusions

What has caused the increase in the price of crude oil after 2003? This highly discussed question has been at the heart of this paper. While competing explanations have been put forward by the academic society, the hypothesis of current demand increases due to strong and unexpected economic growth of emerging economies has been supported most prominently (Hamilton, Kilian, Krugman). This implies that the market must have been constantly shocked by increases in fundamental demand without being capable of adjusting expectations over a time period of several years.

The major challenge in empirically assessing the relative contribution of supply, fundamental and expectation-driven demand consists in finding appropriate time series approximating the three essential components of the price. While this task is comparably straightforward for supply and business-cycle related demand, finding an appropriate proxy for forward-looking demand has remained a rather unsolved issue in the empirical literature on oil market modeling: as expectations are not observable, the contribution of forward-looking demand activities to the price formation can not be directly inferred. This paper proposes a new proxy for expectation-driven demand activities for a structural decomposition of the crude oil price after 2003. It consists of a time series of all news items relevant for the crude oil market that have appeared on news tickers of one of the world's largest news providers. As information is at the root of the expectation formation process, we consider this time series as indicative of market expectations held at any point in time. The subsequent structural decomposition shows that forward-looking demand activities have played an important role for the price development. Accordingly, shocks from news sentiment have contributed to a majority to the price development. This result implies that the market has been adjusting to expected future market conditions. Thus, we do not find evidence to support the view that unexpected shocks from current demand have driven the crude oil price after 2003.

As this result stands in contrast to the reference literature (Hamilton, Kilian, Krugman), we provide an extended discussion about possible factors driving the result. First, we find that most commonly used time series in empirical assessments of the crude oil market as well as the corresponding empirical model exhibit a structural break in 2003 which most studies have not accounted for, so far. We can show that accounting for such instabilities in the time series have a decisive effect on the estimation results: A re-estimation of Kilian (2009) for the structural break free time period from 2003-2010 yield results in line with ours. The second part of the discussion illustrates the robustness of our results to the choice of the fundamental demand proxy. Last, we investigate whether we can find empirical support for the commonly held view that demand from emerging economies has contributed most to the price development. Through appropriate choices of fundamental demand estimators, we can separate between fundamental demand effects arising from OECD countries and those arising from OECD countries plus major emerging economies such as China and India. Results reveal there is no systematic fundamental demand effect attributable to emerging economies. Thus, this paper concludes that expectation-based demand activities, rather than business-cycle related demand activities have majorly contributed to the price rise. Or, in other words, the price development reflects expected future market conditions rather than unexpected shocks to current market conditions.

2.5 Appendix

2.5.1 Re-Estimation of 3-Variables SVAR

We re-estimate Kilian (2009) for the sub-period of 2003-2010.[24] The VAR model is based on monthly data for

$$\mathbf{y_t} = (prod_t, econact_t, price_t)'$$

where $prod_t$ is the percentage change in global crude oil production, $econact_t$ refers to the economic activity index and $price_t$ is the real price of crude oil. The VAR representation is

$$\mathbf{y_t} = \sum_{i=1}^{9} \mathbf{A_i y_{t-i}} + \mathbf{e_t}.^{25} \tag{2.5}$$

The underlying SVAR allows to model the contemporaneous effects between the variables $\mathbf{y_t}$:

$$\mathbf{A_0 y_t} = \sum_{i=1}^{9} \mathbf{A_i^* y_{t-i}} + \varepsilon_\mathbf{t} \tag{2.6}$$

with $\mathbf{A_i} = \mathbf{A_0^{-1} A_i^*}$ and $\mathbf{e_t} = \mathbf{A_0^{-1}} \varepsilon_\mathbf{t}$. We impose the restriction matrix as in Kilian (2009) as

$$\mathbf{e_t} = \begin{pmatrix} e_t^{\Delta prod} \\ e_t^{econact} \\ e_t^{price} \end{pmatrix} = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{pmatrix} \varepsilon_t^{\text{flow supply shock}} \\ \varepsilon_t^{\text{flow demand shock}} \\ \varepsilon_t^{\text{residual shock}} \end{pmatrix} \tag{2.7}$$

As in Kilian (2009), we use monthly percentage changes of global crude oil production taken from the Energy Information Administration (EIA) as measure of crude oil supply. The refiner acquisition cost of imported crude oil, deflated by the US CPI, is used as proxy for the real price of oil. While Kilian

[24] For a more detailed description of the model see Section **??**.

(2009) uses a self-composed shipping index based on single cargo freight rates provided by Drewry's, the follow-up paper by Kilian and Murphy (2010) use the Baltic Dry Exchange Shipping index which is "essentially identical" (Kilian and Murphy, pg. 6) to the index used in Kilian (2009). As the latter is readily available on data providing platforms, such as Datastream, we also use it here. The shipping index appears to be non-stationary in levels and is thus investigated in growth rates.[26] As in Kilian (2009), we use the the refiner acquisition cost of imported crude oil, deflated by the US CPI, as proxy for the real price of oil. It is expressed in logs. Our data start in February 2003 and range until February 2010.

Figure 2.7 displays the impulse response functions on the price of crude oil for the re-estimated model of Kilian (2009).[27] Neither a flow supply shock nor a flow demand shock lead to a significant increase in the price of crude oil. We find significant effects in the autoregressive part in the price of crude oil.

Figure 2.8 displays the historical decomposition of the crude oil price according to this three-variable model. As to be expected from the impulse response functions, the main driver of the price development seems to come from the residual which is interpreted as precautionary demand in Kilian (2009). Neither cumulative effects from flow supply nor from flow demand contribute in a visible way to the development of the crude oil price. This result stands in contrast to Kilian (2009) and illustrates that the results are sensitive to the selection of the sample period.

[26] Note that Kilian (2009) uses a different operation in order to make the series stationary. The series is detrended and expressed in deviations from trend. Both manipulations yield the same results.

[27] The bootstrap-confidence-interval from price to price appears to be biased. According to Philips and Spencer (2010) this bias is due to the bootstrap OLS estimate of the error covariance matrix in the reduced form VAR which is biased downwards.

Figure 2.7: Impulse response function of the price for re-estimated Kilian Model

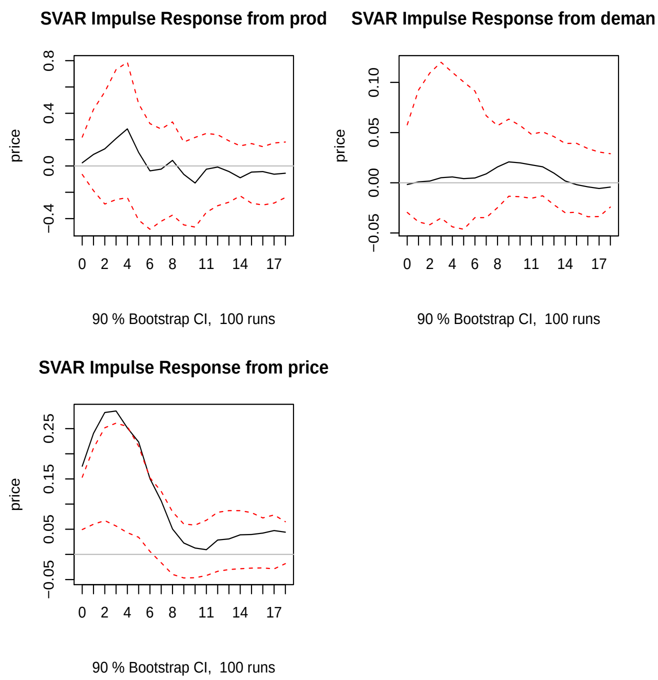
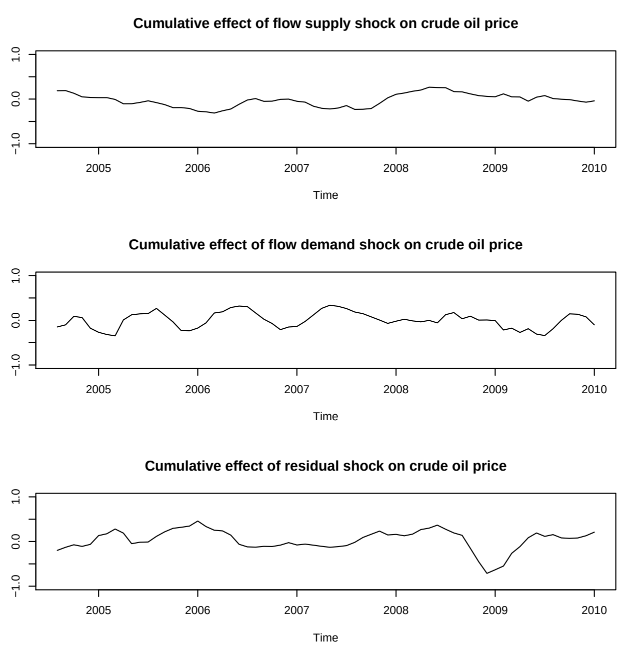


Figure 2.8: Decomposition of crude oil price in three variables model



2.5.2 Re-Estimation of 3-Variables SVAR with OECD Production Indicator

Figure 2.9: Impulse response function of crude oil price with alternative aggregate demand measure

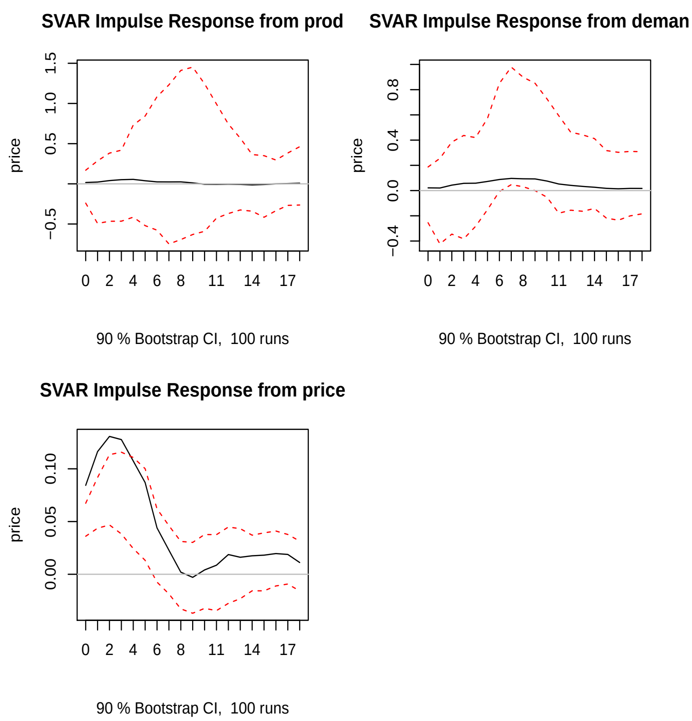
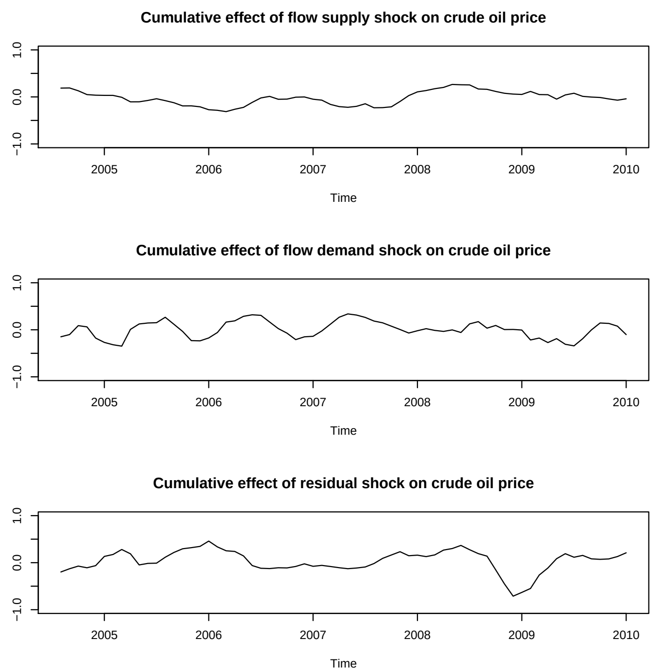


Figure 2.10: Decomposition of crude oil price with alternative aggregate demand measure



2.5.3 Test for Unit Roots

Table 2.3: Test for unit roots / stationarity tests

	ADF test (k=4)	ADF test (k=3)	PP test	KPPS test
Crude oil price	non stationary	non stationary	non stationary	non stationary
Crude oil production	stationary	stationary	stationary	non stationary
Shipping index	non stationary	non stationary	non stationary	non stationary
OECD production	non stationary	non stationary	non stationary	non stationary
Media sentiment	stationary	stationary	stationary	non stationary
CLI	non stationary*	stationary	non stationary	non stationary

*stationary at 15 % level

ADF test: Augmented Dickey-Fuller test, see Dickey and Fuller (1981)

PP test: Philips-Perron test, see Philips and Perron (1988)

KPPS test: see Kwiatkowski *et al.* (1992)

2.5.4 Impulse Response Function of 4-Variable System

Figure 2.11: Impulse response function for supply

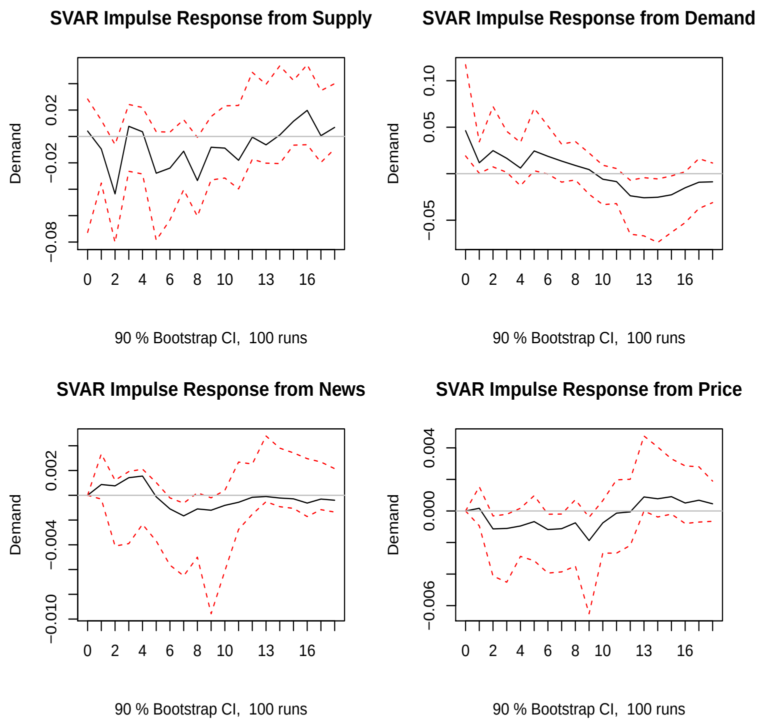


Figure 2.12: Impulse response function for demand

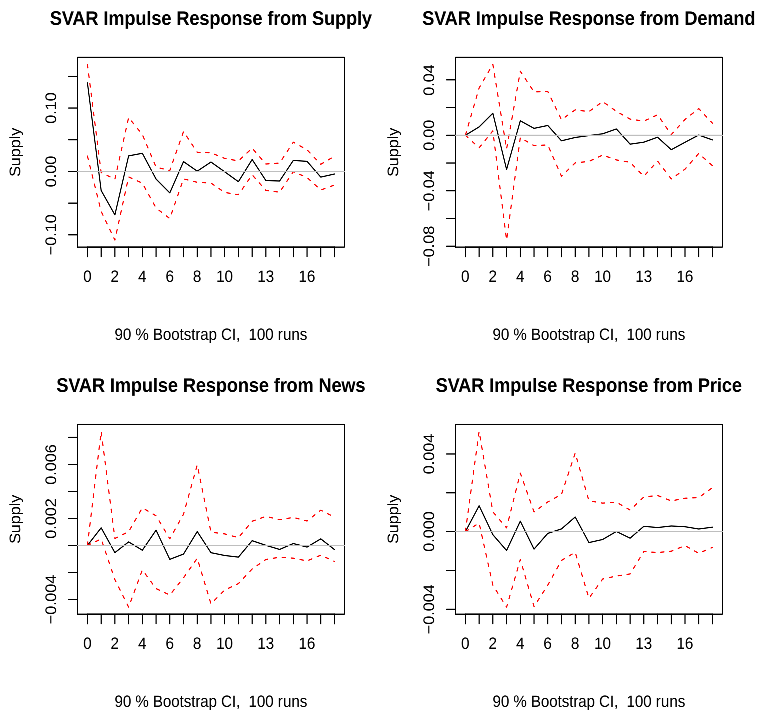
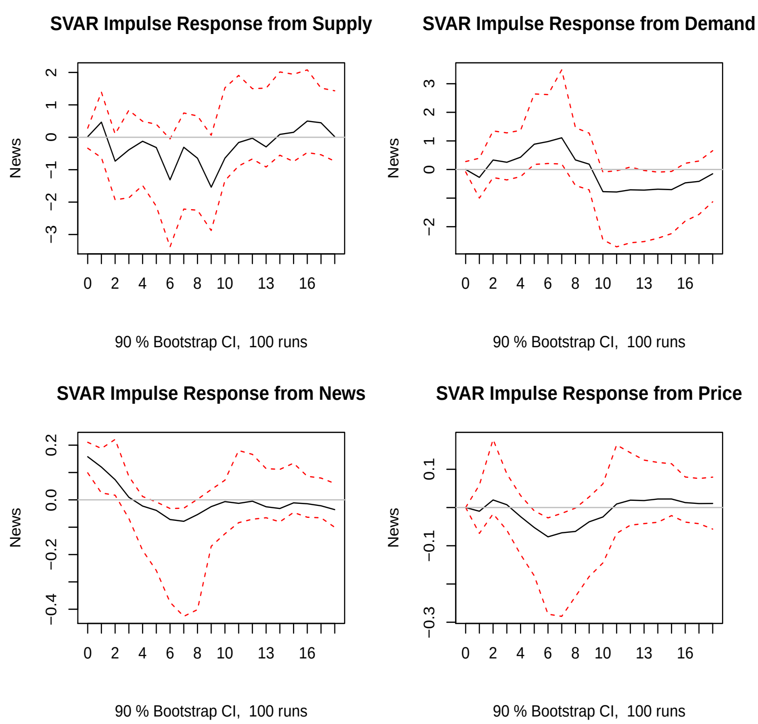


Figure 2.13: Impulse response function for news



2.5.5 Diagram of Fit

Figure 2.14: Diagram of fit and residual for supply

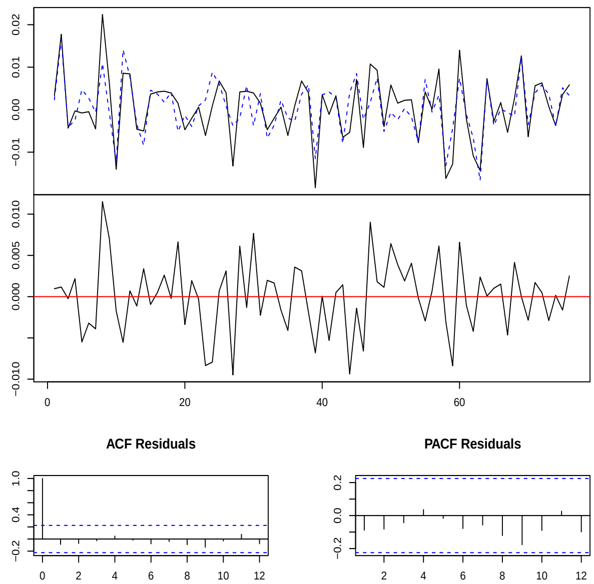


Figure 2.15: Diagram of fit and residual for demand

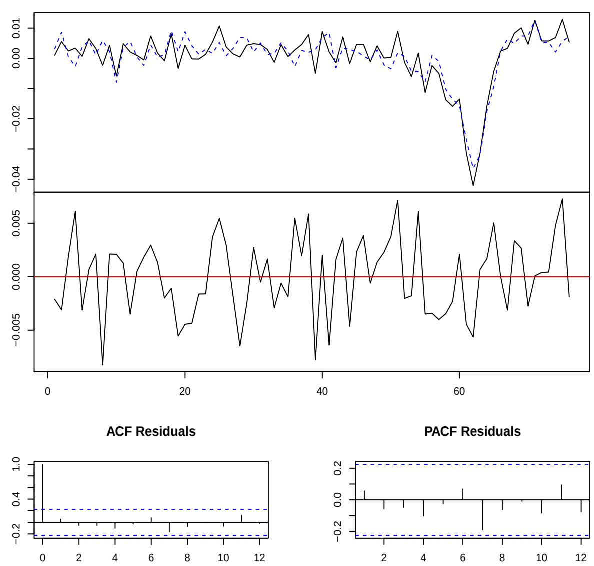


Figure 2.16: Diagram of fit and residual for news

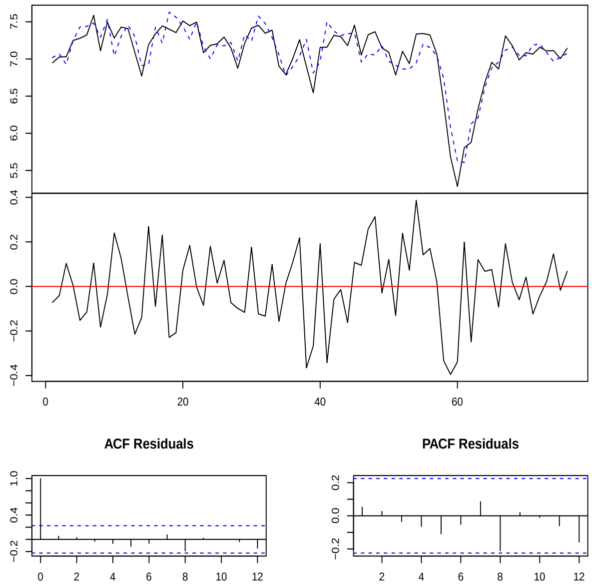
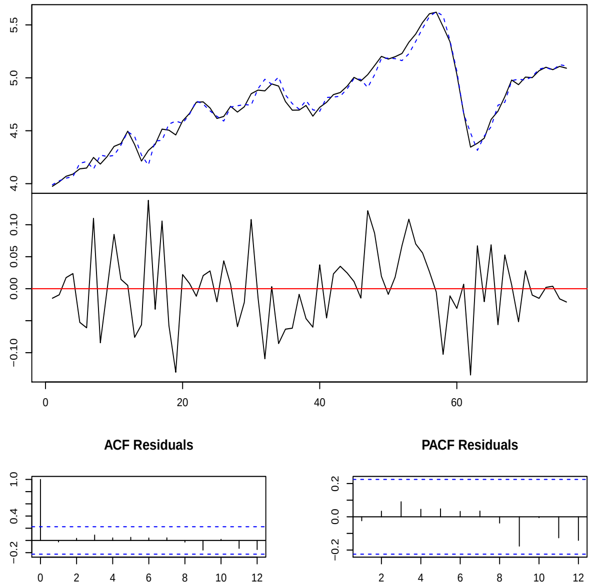


Figure 2.17: Diagram of fit and residual for price



2.5.6 Correlations

Table 2.4: Correlations

	Shipping index	OECD prod	CLI	Petr. cons	Net sent	Cumul. sent	Oil price
Shipping index	1.000	-0.028	0.210	0.045	0.226*	-0.067	-0-052
OECD prod	-	1.000	0.648**	0.381**	0.195	0.765**	0.644**
CLI	-	-	1.000	0.522**	0.531**	0.160	0.177
Petr. cons	-	-	-	1.000	0.217*	-0-056	-0.256*
Net sent	-	-	-	-	1.000	-0-031	0.086
Cumul. sent	-	-	-	-	-	1.000	0.815**
Oil price	-	-	-	-	-	-	1.000

Chapter 3

Energy Use and Economic Growth: Empirical Evidence from Cross-Country Analysis

This paper empirically investigates the causal effect of energy use on a country's economic growth throughout different stages of development. Along with direct effects, energy is allowed to influence income growth indirectly by capital accumulation through input substitution. The crucial findings are that energy use affects economic growth primarily through the capital channel and that this result varies substantially with regard to a country's income level. For high income countries, a higher energy input tends to reduce capital accumulation harming economic growth indirectly, whereas for middle income countries an increase in energy use drives capital accumulation, which in turn pushes economic growth. No significant results are found for low income countries.

3.1 Introduction

Nowadays, increasing global energy demand coming mainly from heating, electricity and transport fuels is primarily met by energy supply based on combustion of fossil fuel. According to the International Energy Agency (IEA) (2006), fossil fuel will account for 77% of the increase in world primary energy demand between 2007 and 2030. Current CO_2 emissions caused by global energy use account for about 80% of total emissions, making energy use a major source of pollution. Hence, possible implications of energy use with regard to sustainable development have become popular topics in public and scientific debates. In this respect, the relationship between energy use and economic growth deserves closer attention: if, as it is widely believed, a stricter energy supply hampers a country's income growth, policy measures targeting a lower energy consumption may have adverse effects on economic development. How does economic research explain this issue? Although a lot of work has been conducted on the role of energy use at the micro level, the importance of energy as causal factor in economic development at an economy-wide level has not been unambiguously worked out so far. Contrary to prevalent views, by modeling a multi-sectoral economy, Bretschger (2010) suggests that a lower energy input can enhance growth through capital accumulation. Intuitively, if a country lowers its energy input due to higher energy prices, it may, under certain conditions, release labor from energy-intensive production sectors to less energy-intensive capital-producing sectors, which in turn increases economic growth.

The empirical evidence on the causal relationship between energy use and income is frequently and hotly debated in applied economic analysis. Results vary with regard to different countries, research periods and econometric methods. Furthermore, different indicators have been used to represent economic output such as income in levels, income growth, and employment rates. Lee (2006) and Huang et al. (2008) provide a systematic survey on econometric studies conducted in this field. To date, in most of these empirical studies, time series methodology is applied, identifying short-run responses and examining the bi-directional relationship applying Granger Causality tests. If uni-directional causality runs from energy use to income, the economy is expected to be energy-dependent. Hence, energy conservation policies are expected to harm economic growth. Accordingly, if causality runs from income to energy use, conservation policies may have few adeverse income effects. Kraft and Kraft's (1978) contribution is frequently considered as pioneering work in this field. Using US data from 1947 to 1974 they find that causality runs from income to energy consumption. Using monthly data for the US, Akarca and Long (1979) find evidence for the opposite effect. In the subsequent years more studies followed which showed mixed results. Akarca and Long (1980), Yu et al. (1988) and Yu and Jin (1992) provide evidence for a neutral relation between energy use and income, whereas Erol and Yu (1987) show that causality runs from energy to income for Japanese data. With the advance of statistical methodology in time series econometrics, the topic has been revised in more recent years. Cheng and Lai (1997), Glasure and Lee (1997) and Soytas and Sari (2003) use Engle-Granger's cointegration test and the error correction mechanism in order to distinguish between short- and long-run effects. Results are mixed and do not show any consistent pattern in the interrelation between income and energy use. More recent studies focus on multivariate Granger causality analysis including more variables in the statistical procedures. Oh and Lee (2004) as well as Paul and Bhattacharya (2004) use several variables from the demand and supply side to analyse the effect of energy use on GDP in South Korea and India respectively, where GPD is found to led energy consumption. Lee and Chang (2007) and Huang et al. (2008) use Panel VAR models estimating parameters over a larger sample of countries. Results remain mixed with energy use exhibiting ambiguous effects on economic output depending on the analyzed time period and country sample. Other than exising time series analysis of higher frequency data, Bretschger (2010) uses an econometric approach related to the growth empirics literature[1] in order to analyse the long-run reaction of an economy in the absence of short-run business-cycle fluctuations. Using a sample of high income countries, he finds that a reduction energy use can enhance economic growth by fostering the capital accumulation process.

The aim of this study is to examine empirically how energy use per capita affects a country's economic growth. We extend Bretschger's (2010) work by analysing a broad set of countries including low- and middle-income country groups accounting explicitly for the role of economic development. We use a dynamic model for panel data, focusing on cross-country variation in order to identify the long-run impact of energy consumption on economic growth. Additionally, we consider possible indirect effects transmitted through the capital accumulation channel caused by possible input substitution effects.

Results suggest that growth rates in high and low income countries are barely affected by higher energy consumption, whereas we find clear evidence that income growth in middle income countries does rely on energy use as a production factor. Accordingly, economic consequences of energy reducing policy measures vary depending on a country's stage of development.

The remainder of the paper is organized as follows. Section 3.2 gives a basic conceptual framework on the interrelation between energy use and economic groth and provides stylized facts. Section 3.3 describes the empirical strategy and section 3.4 presents the empirical results. Section 3.5 summarizes and concludes.

[1] See Mankiw et al. (1992).

3.2 Theoretical Considerations and Stylized Facts

Theoretical basis

In this section we summerize the basic idea of the theoretical framework developed by Bretschger (2010), which analyses the interdependence between growth dynamics and energy use. As in his approach we use the core equations of a simplified growth model in order to illustratively disentangle possible direct and indirect effects of energy use on economic growth. The main insights will be used to derive empirically testable hypothesis.

In economic theory energy (E) is regarded as an input factor for the production of output Y such as capital (K) or labor (L)

$$Y = F(E, K, L). \tag{3.1}$$

In a simplified setting, $F(\cdot)$ is assumed to be a static linear homogenous production function with $\alpha, \beta, \gamma > 0$,

$$Y = K^{\alpha} L^{\beta} E^{\gamma}. \tag{3.2}$$

Since $\frac{\partial Y}{\partial E} > 0$, an increase of energy will lead *ceteris paribus* to an increase in Y. Hence, the more energy is feeded into production activity, the higher will be the corresponding output. In order to analyze the growth process we consider equation (3.2) as a dynamic production function. Taking the derivatives with respect to time and calculating growth rates leads to the well-known growth accounting relation where the hats denote the growth rates:

$$\hat{Y} = \alpha \hat{K} + \beta \hat{L} + \gamma \hat{E}. \tag{3.3}$$

We can see that output growth is driven by the growth rates in inputs. Again, a higher energy input in terms of a higher growth rate favors economic growth. However, this simple framework does not suffice for an appropriate identification of causal effects. Of course, growth accounting is only concerned with the immediate determinants of growth while ignoring essential economic aspects, such as possible interrelations between the inputs as well as bidirectional effects between output and inputs.

In order to analyze the effect of energy use on economic performance, we consider a multi-sectoral economy. Final output is produced with capital and an intermediate input flow, where labor and energy are used as basic inputs to produce capital and the input flow X. Thus, energy and labor can be employed either in the capital sector (L_K, E_K) or in the intermediate good sector (L_X, E_X). The setting of the general model is,

$$Y = F(K, X), \tag{3.4}$$

$$\dot{K} = G_K(L_K, E_K)\kappa - \delta K, \tag{3.5}$$

$$X = G_X(L_X, E_X), \tag{3.6}$$

with the dot denoting the time derivative, and δ being the depreciation rate $(0 < \delta < 1)$. Thus, the differential equation (3.5) represents the formation of capital over time. Based on new growth theory, as proposed by Romer (1990), κ represents positive learning spill-overs from past investments, i.e. $\kappa = K^\eta$ with $0 < \eta < 1$. Like $F(\cdot)$, $G(\cdot)$ represents a linear homogenous production function which is increasing with its inputs L_K and E_K. Total labour supply is given by $L = L_X + L_K$ and total energy supply by $E = E_X + E_L$. For the following argument, we consider the production in the capital sector to be of low energy intensity relative to the intermediate sector X. For simplicity, we assume $G_K(\cdot)$ to be totally independent of energy, i.e. $E = E_X$. Equation (3.5) can be rewritten as follows:

$$\dot{K} = G_K(L_K)\kappa - \delta K. \tag{3.7}$$

Hence, a reduction of energy supply $\Delta E < 0$ implies a reduction of energy input in in the production of good X. Assuming additionally that L_X and

E_X exhibit a low substitutability, a lower energy supply will release labor from sector X towards the capital producing sector, which increases capital production G_K and therewith capital accumulation $\dot{K}$. Thus, under the conditions that elasticities of substitution in sector X are low and that sector K has a low energy intensity, the reduction of total energy supply releases labor into the capital sector which favors capital accumulation leading finally to positive growth effects. The question whether these conditions apply, requires empirical investigations. Depending on data availability, we can test directly whether the properties related to energy-intensity and elasticity of substitution are fullfilled and we can evaluate how energy use is related to the capital accumulation and the growth process. In this paper we will follow the latter approach.

With regard to the empirical specification, the link between capital accumulation and economic growth can be illustrated by considering the conditional convergence property discussed in Mankiw et al. (1992). Accordingly, based on the characteristics of the Solow model, countries reach different steady states of income per capita depending on the determinants of the steady state. Considering equation (3.5), the long run equilibrium value for capital can be derived,

$$K^* = (\frac{G_K(L_K)}{\delta})^{\frac{1}{1-\eta}} \tag{3.8}$$

which positively depends on $G(\cdot)$. Under the conditions assumed above, countries with a lower/higher energy input will have a higher/lower steady state capital value and, consequently, a higher/lower steady state value for income Y^*. Using a first-order Taylor-series approximation the behaviour of the system around the long-run equilibrium can be studied,

$$\frac{d \ln Y}{dt} = \lambda[\ln Y - \ln Y^*], \tag{3.9}$$

with λ being the speed of convergence. Following Mankiw et al. (1992), the growth equation for estimation which is described in the next section is directly derived from term (3.9), with the growth rate of income depending positively on the predicted steady state value and negatively on the actual level of income.

Of course, this indirect interrelation between energy use and growth in a multi-sectoral economy reflects a long-run structural relationship, where the economy needs sufficient time to reallocate input factors. Accordingly, we can think of the simpler approach represented by the function (3.2) or (3.3) as a short run relation, reflecting the immediate reaction on changes in energy input, such as business-cycle effects. In the empirical analysis, we will allow for direct effects of energy use on economic growth as well as for indirect effects transmitted through capital accumulation as indicated by equation (3.5).

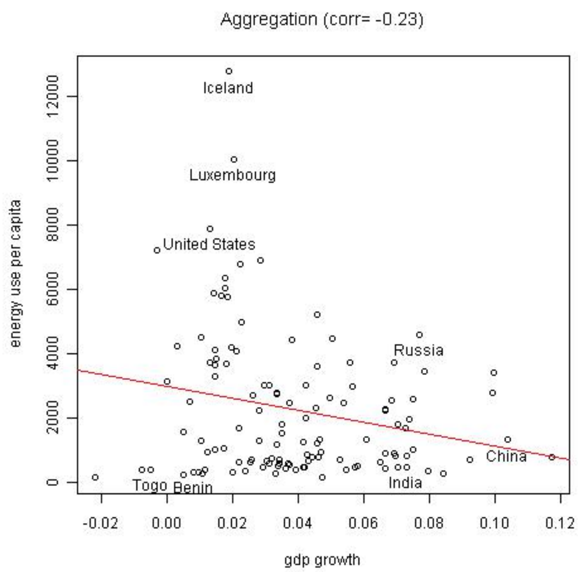
For both theoretical considerations above, the direct and the indirect specification, energy input is assumed to be given exogenously. That is energy use is not determined within the model, which implicitly presumes an inelastic energy supply being determined by natural factors only. Of course, this implication does not reflect the real situation. Countries do influence energy consumption through the price by imposing taxes. Furthermore, considering energy to be closely related to non-renewable resources, economic theory suggests energy prices as well as the optimal production path to be related to the interest rate (Hotelling, 1931). Therefore, energy use is expected to be determined within its macroeconomic context, with aggregated demand Y exhibiting a reversed effect on the energy input. Concerning this matter, we address the statistical identification problem in the empirical section.

Stylized facts

Probably, the most prominent energy-income relations in the recent past were the two oil crisis in the 1970s. Both events were followed by a worldwide recession, indicating a negative relation between energy and economic output. However, this impression changes if we look at broad range of countries for the last years. As it is visualized in figure 3.1, the linear association between energy use per capita and economic growth for a sample of 134 countries averaged over the years from 2003 to 2008 is negative. That is, without controlling for other characteristics, countries with higher growth rates tend to have a lower energy consumption per capita, which is not surprising if we think about fast-growing transition countries such as China and India having a relatively low energy consumption per capita. If we consider the income groups of countries[2] in figure 3.2 we see that among low- and middle-income countries a high energy use per capita is associated with high economic growth rates, whereas the converse applies to high income countries suggesting that income-growth in lower-income countries rely on energy use more heavily. According to Jemelkova and Toman (2003) energy use per unit of output declines over time in the highest-income countries, which may be explained by the application of more efficient technologies as well as the change of economic activity, indicating a decreasing relevance of energy as a production factor. In fact, the main drivers of global primary energy consumption are emerging markets: 93% of the global increase in primary energy between 2007 and 2030 are expected to come from non-OECD countries, driven mainly by China and India (IAE, 2009). Hence, it is expected that the relation between energy use and economic activity varies along different stages of development. Stated differently, the proposition is: in accordance with the Environmental Kuznets Curve hypothesis indicating an inverted U shape relationship between environmental degradation and income per capita, this descriptive analysis suggests that the relationship between energy consumption and economic growth differs with regard to the living standard of the countries. For countries at a low income level, energy use affects economic growth positively, whereas it turns negative the higher stage of development. An appropriate econometric analysis is required in order to identifiy the causal relation of energy use on the economic performance throughout different groups of income.

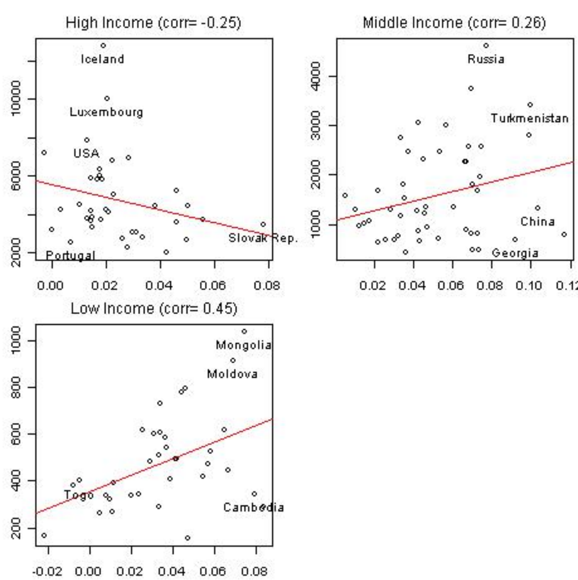
[2]The country selection and grouping will be explained in more detail in the empirical section.

Figure 3.1: Energy use and economic growth (both per capita)



Data Source: World Bank and Penn World Table (PWT 6.3).

Figure 3.2: Energy use and economic growth (both per capita) for country groups



Data Source: World Bank and Penn World Table (PWT 6.3).

3.3 Empirical Methodology

This paper uses a dynamic model for panel data in order to identify the long-run impact of energy consumption on economic growth. The analysis is conducted with a broad sample of countries, thereby accounting for different stages of economic development. As described in the previous section, we consider possible indirect effects transmitted through the capital accumulation channel. The data will be averaged over 5 years in order to exclude short-run business cycle effects. Following Mankiw's et al. (1992) seminal contribution in growth empirics, the growth equations are derived from a linearization of the system around the steady state (see equation (3.9)). Appropriate econometric techniques are applied in order to adequately adress typical estimation problems related to growth regression models.

3.3.1 Growth Regressions

The methodological debate on the statistical identification of the determinants of economic growth has a long lasting tradition. Basically, the major drawbacks include "an often excessive distance between measured variables and the theoretical concepts they are meant to capture; poor grounding of estimated functional forms in economic theory (...) and a small number of available observations (Hauk and Wacziarg 2009, p.104)" giving rise to several estimation difficulties such as endogeneity, persistent variables, measurement errors, omitted variables and parameter heterogeneity. Corresponding to the theoretical consideration above, previous empirical studies on energy consumption and economic growth show that the presence of endogeneity due to a possible feedback relationship (simultaneity) is indeed a major concern. In the following we present the econometric strategy used to identify the causal effect of energy use on economic growth.

Following the usual notation in the econometric literature the growth regression model for panel data can be written as:

$$\Delta y_{it} = (\gamma - 1)y_{i,t-1} + x_{it}'\beta + \alpha e_{it} + \eta_t + \mu_i + \epsilon_{it}. \tag{3.10}$$

y_{it} is log per capita income at period t in country i and Δy_{it} is the corresponding first difference representing the growth rate of income per capita during period t. x_{it} is a vector of control variables varying over countries and over time. The variables are (log) investment rates, population growth rates, a measurements for human capital (approximated by the gross primary school enrollment rate) and trade openness. η_t and μ_i are period-specific and country-specific variables respectively. In order to control for possible direct effects of energy use on economic growth we include log energy consumption per capita e_{it} as a regressor in the estimation equation (3.10). The functional form of the estimation equation is derived from the Solow growth model, which is the prevalent theoretical fundament in the empirical growth literature.

Compared to OLS, in the absence of possible sources of endogeneity, one can achieve more efficient estimates by applying a conventional panel-technique such as the Fixed effects (FE) or the Random effects (RE) method to equation (3.10), which are commonly used for static model specifications. Assuming the unobserved individual-specific time-invariant variables μ_i, such as the initial level of technology, to be correlated with included right-hand side variables, the FE-estimator will dominate the RE-estimator because it prevents bias caused by omitting this unobserved heterogeneity. We perform Hausman specification tests based on Hausman (1978) in order to detect failures in the assumptions for the RE model.

However, the methods mentioned above usually assume strict exogeneity for all right-hand-side variables, that is

$$E(e_{i,t}|X_i) = 0, t = 1, ..., T,$$

where X_i is a vector including all regressors. Obviously, this assumption cannot be maintained in the presence of a lagged dependent regressor $y_{i,t-1}$.[3] Additionally, as described in the previous section, it is expected that energy consumption and economic growth are determined simultaneously, which again violates the strict exogeneity assumption. Arellano and Bond (1991) propose a GMM estimator in first differences using sequential exogeneity only for consistent parameter estimation. This method is briefly described in the following.

According to Bond et al. (2001), equation (3.10) can be rearranged as follows

$$y_{it} = \gamma y_{i,t-1} + x'_{it}\beta + \alpha e_{it} + \eta_t + \mu_i + \epsilon_{it}. \tag{3.11}$$

Taking first differences:

$$\Delta y_{it} = \eta_t - \eta_{t-1} + \gamma \Delta y_{i,t-1} + \Delta x'_{it}\beta + \alpha \Delta e_{it} + \Delta \epsilon_{it}. \tag{3.12}$$

Under suitable assumptions, valid instruments can be found for all regressors included in equation (3.12).[4] The corresponing orthagonality condition in the case of energy consumption is:

$$E(e_{i,t-\tau}\Delta\epsilon_{it}) = 0, t = 2, ..., \tau; \tau \geq 2.$$

Blundell and Bond (1998) show that in the presence of persistent data due to weak instruments, the GMM estimator for the model in first differences is poorly behaved in finite samples. They propose supplementary moment conditions based on additional model equations in levels leading to the so-called System-GMM estimator.

In correspondence with the argumentation above, Bond et al. (2001) suggest that, theoretically, the GMM and the SYS-GMM methods should be prefered to more conventional estimators when dealing with growth regressions

[3]The bias arising in dynamic panel data models due to the lagged dependent variable is treated in detail in chapter 5.

[4]For more detailed information see Arellano and Bond (1991).

due to consistency reasons. However, considering a finite sample, the estimation properties strongly depend on the data generating processes. Based on Monte Carlo simulations Hauk and Wacziarg (2009) evaluate bias properties of different econometric methods which are frequently used in the growth regression context. Basically, they show that no method can be applied to overcome all possible sources of bias simultanously. We obtain similar results in the simulation study in chapter 5. For example, by adressing the reverse causality bias with the GMM technique, the bias coming from measurement error increases relative to the other methods. In order to check if results are robust, we therefore apply different estimation methods such as pooled OLS, RE, FE, GMM and SYS-GMM. Furthermore, the performance of the GMM estimators strongly depends on the validity of the instruments. Accordingly, we use the Hanson's test of overidentifying restrictions based on Hanson (1982) in order to test whether the restrictions of the GMM-models are satisfied. A rejection of the test indicates a possible failure of the moment conditions. We additionally test the subset of orthagonality conditions included for the SYS-GMM estimation by the so-called Difference-in-Sargan test developed by Eichenbaum et al. (1988). The rejection of the null indicates possible invalidity of the additional instruments and thus a failure of the SYS-GMM assumptions.

3.3.2 Capital Regressions

Based on the theoretical considerations above, additional regressions are performed examining the effect of energy consumption on capital accumulation, representing a possible channel through which economic growth can indirectly be influenced. Similar estimation techniques are applied as for the growth regressions, expecting capital accumulation and energy use to be simultaneously determined. Apart from energy use, the control variables included reflect the demographic and economic structure of the corresponding country. We consider two empirical specifications. The first is static, where ci is the average investment share of real GDP, approximating the accumulation of physical capital,

$$ci_{it} = z'_{it}\gamma + \delta e_{it} + \eta_t + \mu_i + \epsilon_{it}. \tag{3.13}$$

Again, e_{it} is log energy use per capita, η_t and μ_i are period-specific and country-specific variables respectively and z_{it} is a vector of control variables including population, population growth, the government share of real GDP, trade openess, the ratio of the dependent population (younger than 15 or older than 65) to the working age population, the share of urban population and life expectancy at birth. The second specification of the capital regression is dynamic, as it includes the lagged dependent variable $ci_{i,t-1}$. Due to consistency reasons, the latter version is estimated only by GMM and SYS-GMM methods. Similar tests are performed as for the growth regression estimations.

Throughout all estimations, robust standard errors are applied controlling for hetereoscedasticity and autocorrelation in the error term.

3.3.3 Data

The empirical analysis includes 117 countries, which are subdivided into 3 income groups using GNI for the year 2008 as proposed by the World Bank. As it is shown in table 3.1, the high income group includes 36 countries, the middle income group includes 47 countries, and the low income group includes 34 countries.[5] A description of the variables as well as data origin is given in table 3.2. The time period ranges from 1973 to 2007. Taking 5-years

[5]Exclusion of countries is mainly due to poor data availability. Moreover, being highly incomplete, energy prices could not be included as a control variable. Likewise, due to data incompleteness regarding school enrollment rates, growth regressions are estimated for the full sample without regressing on school enrollment as well as with a reduced sample including school enrollment rate as a regressor.

averages we obtain 7 observations over time.

Table 3.1: Countries included in the analysis

High income group
Australia, Austria, Belgium, Brunei, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea Republic, Luxembourg, Malta, Netherland, New Zealand, Norway, Oman, Portugal Saudia Arabia, Singapore, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, the UK, USA
Middle income group
Albania, Algeria, Argentina, Armenia, Belarus, Bosnia and Herzegovina, Botwana, Brazil Bulgaria, Chile, China, Colombia, Costa Rica, Domenican Republic, El Salvador, Gabon Georgia, Guatemala, India, Iran, Jamaica, Jordan, Kazakhstan, Latvia, Lebanon, Libya Lithuania, Macedonia, Malaysia, Mexico, Morocco, Namibia, Panama, Paraguay, Peru Poland, Romania, Russia, Serbia, South Africa, Syrian Arab Republic, Thailand, Tunisia Turkey, Turkmenistan, Ukraine, Uruguay, Venezuela
Low income group
Bangladesh, Benin, Bolivia, Cambodia, Cameroon, Democratic Republic of the Congo Egypt, Eritrea, Ethiopia, Ghana, Haiti, Honduras, Indonesia, Kenya, Kyrgyz Republic Moldova, Mongolia, Mozambique, Nepal, Nicaragua, Nigeria, Pakistan, the Philippines The Republic of Congo, Senegal, Sri Lanka, Sudan, Tajikistan, Tanzania, Togo Togo, Vietnam, Yemen, Zambia

Table 3.2: Variables included in the analysis

Variable	Description	Origin
y	real per capita GDP, constant prices, chain series	PWT 6.3
g	real per capita GDP growth, constant prices, chain series	PWT 6.3
enuse	energy use per capita (kg of oil equivalent per capita)	World Bank
ci	average investment share	PWT 6.3
pop	population	PWT 6.3
popg	population growth	PWT 6.3
enroll	school enrollment primary (% gross)	World Bank
openc	(exports + imports)/GDP	PWT 6.3
cg	governmet share of real GDP	PWT 6.3
agedep	age dependency ratio ((¡15)+(¿65))/working age pop.)	World Bank
shurb	share of urban population	World Bank
exp	life expectancy at birth	World Bank

3.4 Empirical Evidence

3.4.1 Estimation Setting

The results presented in this section cover the estimations for the growth and capital regression. Each regression is applied to the aggregate sample of countries, as well as to the three different income groups of countries in order to control for parameter differences with regard to different stages of development. Along with the GMM and SYS-GMM approach, I apply pooled OLS as well as the RE and FE estimators. The estimated coefficients and the different test results are listed in the following tables, which can be found in the appendix: table 3.4 and table 3.5 show the growth regressions for the aggregated sample as well as for each income group.[6] Table 3.6 to table 3.9

[6] Each regression is performed twice: once with a proxy for human capital accumulation (school enrollment rate) and once without. As mentioned before, in order to include the school enrollment variable, I had to reduce the corresponding sample size.

show the estimated coefficients for the static capital regressions, whereas in table 3.10 we can find the results for the dynamic specifications. Some additional test results are listed in table 3.11. Table 3.3 represents a clearly arranged overview over all estimation results regarding the causal effect of energy use on economic growth. Energy use in the growth regression reflects the direct effect on economic growth, summerized in the upper part of this table, whereas energy use in the capital regression shows a possible indirect effect listed in the lower part of this table.

3.4.2 Growth Regression Results

Table 3.4 and table 3.5 show that for the major part of the estimations, energy use does not affect economic growth directly, neither in the aggregated case nor for the single groups. A significant direct effect is only found using SYS-GMM for high-income countries (column (20) in table 3.4). Thus, by applying growth regressions on different income groups of countries, I find energy neutrality in most cases. In the following, I will describe the estimation results in more detail.

Considering the aggregated sample, we can see in table 3.4 that the estimated coefficients generally correspond with their expected values from theoretical and empirical growth literature.[7] The coefficient for last period income has a negative and significant effect on GDP growth in every regression, reflecting conditional convergence. The growth enhancing effect of capital accumulation is reflected in the significant and positive effect of investment share for physical capital throughout most regressions. Expectedly, population growth has a significant negative and trade openness a significant positive effect, indicating growth advantages for more open economies. When including school enrollment rate, the results for GMM and SYS-GMM do not appear to be robust with respect to the convergence rate. The expected significant posi-

[7] See Mankiw et al. (1992) or Islam (1995) for seminal contributions.

tive effect for school enrollment rate can only be found in column (15). In accordance with Hauk and Wacziarg (2009) and chapter 5 the GMM estimator tends to detect higher convergence rates then the other estimators listed here.

The estimations for high income countries again show their expected signs and magnitudes. Compared to the aggregated sample, population growth do not seem to play a relevant role at a higher stage of development. The inclusion of school enrollment rates leads to less significant coefficients, especially for the FE estimations. As an exception to all growth regressions performed here, energy use exhibits a positive and significant effect for the SYS-GMM estimator in column (20).

The results for the middle income countries resemble the results for the aggregated country groups with population growth exhibiting a negative effect on economic growth. Compared to the other samples, the results are more robust to the inclusion of the enrollment variable.

The estimation for low income countries tend to exhibit less significant results. However, where significance is obtained, the signs and magnitude of the coefficients reflect the expected value, especially with respect to capital accumulation and population growth. A possible reason for the less concise results in this group might be worse data quality and availability for lower income countries.

In general, with the inclusion of school enrollment rate as a proxy for human capital the explanatory power of all model results changes for the worse. As described in data section above, the sample size has to be reduced in order to account for human capital, which might be a reason for this. In fact, we can see from the usual test results such as the R-squared, the Wald and the Sargan statistic that the goodness and reliability of the models is notably reduced revealing a preference for the estimation equations without school enrollment rate. Nonetheless, throughout all regressions, energy neutrality as well as the positive effect of capital accumulation are robust. Validity of the instruments for the GMM estimators can not be rejected according to the

Sargan test and the Sargan difference test in table 3.11. The Hausman test in table 3.11 indicates an endogeneity problem for the RE estimator favouring the FE technique

3.4.3 Capital Regression Results

More revealing are the results related to the capital regressions shown in table 3.6 to 3.10. Given the usual assumptions on the regressions hold true, results suggest that energy use has a causal effect on capital accumulation which varies with regard to a countries stage of development. However, compared to the growth regressions above, the signs and magnitudes of the other coefficients turn out to be less clear in terms of their interpretation and robustness. A possible reason might be the weak theoretical link between economic theory and the functional form of the estimation equations. In the following, results are described in more detail.

Considering the aggregated country sample in table 3.6, I find that, except for the FE estimator, energy use exhibits a positive and significant effect on physical capital accumulation. Additionally, trade openess is also found to signficantly drive capital accumulation. Further results are less robust throughout all estimation techniques: population growth has a positive and significant effect on capital accumulation indicating that economies that are demographically more dynamic tend to accumulate more capital. On the other hand, life expectancy exhibits a negative and significant effect on economic growth. This effect might be caused by the fact that a higher life expectancy implies also a higher part of non-working population which in turn raises intergenerational distribution harming investment activities. The same argument also applies to the age dependency ratio, which is found to be positive and significant only for the OLS estimation.

The estimations for high income countries in table 3.7 show that energy use has a negative and significant effect on capital accumulation. As for the aggregated country sample, age dependency and life expectancy tend to have a negative effect on capital accumulation. The effect for population growth is significant but not robust.

Energy use has a positive and significant effect for capital accumulation for middle income countries. Again, age dependency and life expectancy tend to have negative effects on capital accumulation. Other then in the case of high income countries, trade openess fosters capital accumulation significantly.

As for growth regressions, the low income group exhibits the least robust results. The only consistent result is the positive and significant effect of trade on capital accumulation. The OLS regression shows a positive and significant effect of energy use, however, age dependency and life expectancy are also positive and significant.

The dynamic specification of capital regressions in table 3.10 reveals that the lagged dependent variable is positive significant for all specification and samples indicating persistance in the capital accumulation process. With regard to energy use, we find that capital accumulation is only affected for middle income countries for the GMM specification, where it positively depends on energy use. Except for trade openess, which tends to have a positive and significant effect, the other coefficients do not reveal any robust results in the dynamic specification. As in the case of growth regressions, the Sargan tests generally support the validity of the instrument and, except for the middle income countries, the Hausman test rejects the consistency of the RE model.[8]

[8]In order to increase sample size and improving finite sample behaviour for the single income groups, we performed all estimation including growth and capital regressions for only two income groups; high income and low income countries. With regard to growth regression the results are robust, where energy use does not have any effect on economic growth. The significant effect of energy use on capital regression, however, disappears. The reason for this might be the fact that energy use is specifically vital for the economies of the middle income group including the main emerging markets such as China and India. If these countries are part of either the high or the low income group, this specific capital enhancing effect averages out.

Table 3.3: Summarized results

	Direct effects	
All countries	*no effect*	
High income countries	*no effect**	
Middle income countries	*no effect*	
Low income countries	*no effect*	

	Indirect effects	
	Static	Dynamic
All countries	*positive effect*	*no effect*
High income countries	*negative effect*	*no effect*
Middle income countries	*positive effect*	*positive effect***
Low income countries	*no effect*	*no effect*

Notes: *Except for one estimation equation

**Except for SysGMM

3.4.4 Summarized results

In this empirical analysis, we allow energy to have two possible effects on economic growth. The direct effect is represented in the growth regression. The indirect effect is identified by regressing capital accumulation on energy use, with the former being also an explanatory variable in the growth regressions. Throughout all country groups for almost all estimation strategies capital accumulation exhibits a positive effect on economic growth. Energy does not show any significant direct effect in most regressions. On the other hand, we find evidence that energy use influences economic growth through the capital channel. This effect is positive for middle income countries. High income countries show a negative indirect effect only when considering the static capital regression. No effects, direct or indirect, can be found for low income countries. The results are summerized in table 3.3.

Thus, results suggest that for high income countries a reduced energy use does not harm economic growth. In fact, some of the model specifications suggest that energy restrictions can even be good for economic growth coming from reallocating inputs toward growth enhancing acitvities such as capital accumulation. In other words, a high energy use can be harmful since capital accumulation is crowded out by abundant energy use which in turn reduces economic growth. Furthermore, contrary to high income countries, empirical evidence suggests that for middle income countries energy use indirectly pushes economic growth through capital accumulation. In accordance with the descriptive facts listed above, countries at a transitional stage of development, i.e. emerging economies, seem to rely on energy use as relvant factor fostering economic growth.

3.5 Conclusions

In this paper we empirically investigate the causal effect of energy use on economic growth. Unlike most previous empirical studies in this field, we focus on the long-run relationship considering a broad range of countries and applying the growth regression approach. Capital accumulation is additionally considered as a possible channel through which economic performance can indirectly be influenced by energy use. It is expected that the link-age among energy use and economic growth varies with regard to the stage of development. Accordingly, we perform the analysis for different income groups seperately. Different estimation strategies such as RE, FE and GMM approaches have been applied in order to consitently identify the effects of interest.

The estimated regressions suggest, that energy does not affect economic growth directly. Furthermore, capital has the expected significant positive effect on economic growth for all country groups throughout most estimation methods. With regard to capital regressions, countries seem to vary in terms of indirect energy effects. For high and low income countries a higher energy input does not harm economic growth directly nor indirectly, whereas for middle income countries, an increase in energy use seems to drive capital accumulation, which in turn pushes economic growth. This confirms that, with regard to economic growth, middle income countries which typically are in a transitional stage of economic development rely more heavily on the use of energy in production, whereas energy reduction is not found to be harmful neither for high nor for low income countries. Possible causes for these differences may lie in economic features which are typical for emerging markets such as the production structure or higher energy intensities. Further research is required to investigate this point. Based on this analysis, we conclude that a country's stage of economic development has to be accounted for when considering the long run consequences for economic growth caused by energy-reducing policy measures.

Table 3.4: Growth regressions: all countries and high income countries

All countries N=732, dep. var.: GDP growth (g)						**High income countries** N=233, dep. var.: GDP growth (g)					
	OLS	RE	FE	GMM	SYS-GMM		OLS	RE	FE	GMM	SYS-GMM
	(1)	(2)	(3)	(4)	(5)		(6)	(7)	(8)	(9)	(10)
y(-1)	-0.016**	-0.021**	-0.082***	-0.108*	- 0.126***	y(-1)	-0.026**	-0.023	-0.070*	- 0.241**	- 0.152***
	(0.008)	(0.009)	(0.025)	(0.062)	(0.048)		(0.014)	(0.015)	(0.038)	(0.105)	(0.039)
enuse	-0.005	-0.004	-0.000	0.004	0.014	enuse	-0.007	-0.011	-0.027	0.150	0.063
	(0.008)	(0.008)	(0.021)	(0.068)	(0.054)		(0.011)	(0.015)	(0.034)	(0.092)	(0.040)
ci	0.035***	0.039***	0.046**	0.187***	0.199***	ci	0.055***	0.055***	0.030	0.242***	0.258***
	(0.008)	(0.009)	(0.021)	(0.043)	(0.037)		(0.013)	(0.013)	(0.036)	(0.066)	(0.048)
popg	-1.095***	-1.148***	-1.034**	-1.684**	–1.834***	popg	-0.057	-0.097	0.297	1.484	1.140
	(0.233)	(0.244)	(0.467)	(0.725)	(0.600)		(0.339)	(0.440)	(1.057)	(1.163)	(0.954)
openc	0.022***	0.026***	0.077***	0.128***	0.135***	openc	0.025***	0.028***	0.129***	0.257***	0.056**
	(0.005)	(0.006)	(0.015)	(0.048)	(0.025)		(0.005)	(0.006)	(0.034)	(0.096)	(0.031)
R-squared	0.146	0.139	0.114				0.271	0.246	0.205		
Wald	<2.2e-16	<2.2e-16	<2.2e-16	<2.2e-16	<2.2e-16		<2.2e-15	<2.2e-13	<2.2e-10	<2.2e-16	<2.2e-16
Sargan				0.008	0.005					0.205	0.039
All countries, with human capital, N=693						**High income countries**, with human capital N=220					
	OLS	RE	FE	GMM	SYS-GMM		OLS	RE	FE	GMM	SYS-GMM
	(11)	(12)	(13)	(14)	(15)		(16)	(17)	(18)	(19)	(20)
y(-1)	-0.013*	-0.016**	-0.068***	-0.032	-0.053	y(-1)	-0.030**	-0.025	-0.023	- 0.222*	-0.190***
	(0.008)	(0.009)	(0.024)	(0.081)	(0.044)		(0.013)	(0.018)	(0.035)	(0.114)	(0.041)
enuse	-0.005	-0.005	-0.008	0.028	-0.041	enuse	-0.005	-0.011	-0.004	0.126	0.099**
	(0.008)	(0.008)	(0.025)	(0.092)	(0.054)		(0.011)	(0.018)	(0.038)	(0.101)	(0.049)
ci	0.033***	0.037***	0.046**	0.227***	0.161***	ci	0.035***	0.060***	0.044	0.231***	0.273***
	(0.008)	(0.009)	(0.022)	(0.038)	(0.037)		(0.015)	(0.017)	(0.032)	(0.066)	(0.053)
popg	-1.033***	-1.053***	-0.850*	-1.215	-2.139***	popg	-0.065	-0.138	-0.172	1.531	1.918
	(0.219)	(0.207)	(0.444)	(1.232)	(0.611)		(0.394)	(0.507)	(1.096)	(1.474)	(1.140)
openc	0.020***	0.024***	0.081***	0.058	0.091***	openc	0.026***	0.029***	0.097***	0.221**	0.071***
	(0.005)	(0.006)	(0.016)	(0.051)	(0.030)		(0.005)	(0.008)	(0.033)	(0.092)	(0.022)
enroll	0.002	0.002	0.001	-0.001	0.014**	enroll	0.001	0.002	0.005*	-0.001	-0.008
	(0.003)	(0.002)	(0.003)	(0.007)	(0.007)		(0.003)	(0.002)	(0.003)	(0.004)	(0.008)
R-squared	0.133	0.127	0.101				0.271	0.256	0.185		
Wald	<2.2e-16	<2.2e-16	<2.2e-14	<2.2e-16	<2.2e-16		<2.2e-14	<2.2e-12	<2.2e-8	<2.2e-16	<2.2e-16
Sargan				0.001	0.007					0.264	0.057

Notes: Robust standard errors in parenthesis

*** p-value $\leq$ 0.01, ** p-value $\leq$ 0.05, * p-value $\leq$ 0.1

Table 3.5: Growth regressions: middle income countries and low income countries

Middle income countries N=280, dep. var.: GDP growth (g)						**Low income countries** N=209, dep. var.: GDP growth (g)					
	OLS	RE	FE	GMM	SYS-GMM		OLS	RE	FE	GMM	SYS-GMM
	(1)	(2)	(3)	(4)	(5)		(6)	(7)	(8)	(9)	(10)
y(-1)	-0.064***	-0.068***	-0.157***	-0.268**	-0.134***	y(-1)	-0.004	-0.015	-0.063**	-0.277**	-0.115
	(0.017)	(0.016)	(0.037)	(0.111)	(0.046)		(0.013)	(0.011)	(0.029)	(0.134)	(0.069)
enuse	0.003	0.002	0.021	0.104	0.042	enuse	-0.018	-0.023	0.018	0.033	-0.171
	(0.014)	(0.012)	(0.030)	(0.074)	(0.051)		(0.014)	(0.016)	(0.034)	(0.136)	(0.116)
ci	0.035**	0.044***	0.125***	0.225***	0.170***	ci	0.024**	0.030**	0.052**	0.138***	0.066
	(0.013)	(0.015)	(0.020)	(0.068)	(0.041)		(0.011)	(0.013)	(0.024)	(0.050)	(0.070)
popg	-1.446***	-1.560***	-2.210***	-3.255***	-1.636*	popg	-1.200***	-1.102**	-0.939	-2.657**	-3.170**
	(0.392)	(0.350)	(0.540)	(1.078)	(0.842)		(0.433)	(0.506)	(0.621)	(1.077)	(1.297)
openc	0.022**	0.026**	0.091***	0.160*	0.121***	openc	0.018	0.032***	0.060***	0.078	0.048
	(0.010)	(0.011)	(0.030)	(0.093)	(0.042)		(0.012)	(0.012)	(0.021)	(0.047)	(0.082)
R-squared	0.210	0.214	0.194				0.080	0.106	0.143		
Wald	<2.2e-10	<2.2e-13	<2.2e-13	<2.2e-16	<2.2e-16		0.001	<2.2e-4	<2.2e-6	<2.2e-16	<2.2e-16
Sargan				0.311	0.526					0.427	0.182
Middle income countries, with human capital, N=240						**Low income countries**, with human capital N=203					
	OLS	RE	FE	GMM	SYS-GMM		OLS	RE	FE	GMM	SYS-GMM
	(11)	(12)	(13)	(14)	(15)		(16)	(17)	(18)	(19)	(20)
y(-1)	-0.057***	-0.058***	-0.139***	-0.148*	-0.034	y(-1)	-0.004	-0.010	-0.081**	- 0.559***	0.104
	(0.017)	(0.014)	(0.043)	(0.088)	(0.016)		(0.013)	(0.012)	(0.035)	(0.180)	(0.085)
enuse	0.004	0.003	0.012	0.043	-0.006	enuse	-0.019	-0.021	0.028	0.078	-0.121
	(0.015)	(0.012)	(0.038)	(0.118)	(0.064)		(0.016)	(0.015)	(0.035)	(0.184)	(0.128)
ci	0.037***	0.042***	0.121***	0.263***	0.093	ci	0.021*	0.023*	0.028	0.179***	0.038
	(0.013)	(0.015)	(0.021)	(0.054)	(0.073)		(0.012)	(0.013)	(0.034)	(0.053)	(0.041)
popg	-1.245***	-1.298***	-1.865***	-3.456***	-3.070***	popg	-1.159**	-1.085**	-0.794	-3.324***	-4.139***
	(0.364)	(0.278)	(0.542)	(1.225)	(0.554)		(0.473)	(0.514)	(0.671)	(1.269)	(0.905)
openc	0.022**	0.025***	0.107***	0.086	0.043	openc	0.014	0.022*	0.069***	0.067	-0.004
	(0.009)	(0.009)	(0.030)	(0.077)	(0.039)		(0.015)	(0.011)	(0.023)	(0.055)	(0.051)
enroll	0.005	0.004	-0.006	0.001	0.018	enroll	-0.001	0.001	0.003	0.010	0.026
	(0.005)	(0.005)	(0.006)	(0.009)	(0.013)		(0.006)	(0.005)	(0.006)	(0.014)	(0.026)
R-squared	0.169	0.176	0.160				0.040	0.064	0.103		
Wald	<2.2e-10	<2.2e-9	<2.2e-8	<2.2e-16	<2.2e-16		0.027	0.034	0.001	<2.2e-7	<2.2e-16
Sargan				0.234	0.328					0.107	0.127

Notes: Robust standard errors in parenthesis

*** p-value $\leq$ 0.01, ** p-value $\leq$ 0.05, * p-value $\leq$ 0.1

Table 3.6: Capital regressions: All countries

All countries N=732, dep. var.: investment share (ci), static					
	OLS	RE	FE	GMM	SYS-GMM
	(1)	(2)	(3)	(4)	(5)
enuse	0.157***	0.182***	0.069	0.682***	0.815***
	(0.030)	(0.065)	(0.076)	(0.231)	(0.161)
pop	-0.001	-0.039	-0.245*	-0.336	0.004
	(0.017)	(0.035)	(0.145)	(0.214)	(0.053)
popg	-1.583	2.291*	2.972**	1.133	1.959
	(1.346)	(1.187)	(1.233)	(1.345)	(2.689)
cg	-0.001	0.001	0.001	0.001	-0.001
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
openc	0.077*	0.167***	0.255***	0.289***	0.027
	(0.043)	(0.051)	(0.056)	(0.075)	(0.084)
agedep	-0.337**	-0.030	-0.115	0.149	0.150
	(0.143)	(0.177)	(0.209)	(0.228)	(0.139)
shurb	0.151**	0.027	0.045	-0.118	-0.687**
	(0.060)	(0.111)	(0.169)	(0.276)	(0.279)
exp	0.071	-0.179***	-0.125*	-0.154*	-0.225
	(0.075)	(0.062)	(0.058)	(0.088)	(0.173)
R-squared	0.337	0.184	0.087		
Wald test	<2.2e-16	<2.2e-16	0.001	<2.2e-16	<2.2e-16
Sargan test				0.001	0.001

Notes: Robust standard errors in parenthesis

*** p-value $\leq$ 0.01, ** p-value $\leq$ 0.05, * p-value $\leq$ 0.1

Table 3.7: Capital regressions: High income countries

High income countries N=240, dep. var.: investment share (ci), static					
	OLS	RE	FE	GMM	SYS-GMM
	(1)	(2)	(3)	(4)	(5)
enuse	-0.164***	-0.242***	-0.318**	-1.039**	-0.095
	(0.035)	(0.091)	(0.136)	(0.451)	(0.236)
pop	-0.026	0.006	0.202	0.948	0.016
	(0.018)	(0.045)	(0.311)	(0.601)	(0.030)
popg	-5.346***	-0.126	3.479*	4.719*	-7.802***
	(1.506)	(1.062)	(1.879)	(2.513)	(1.831)
cg	-0.003*	0.003	0.007	0.002	-0.007*
	(0.001)	(0.006)	(0.007)	(0.010)	(0.003)
openc	-0.050	0.047	0.173*	0.176	0.088
	(0.043)	(0.069)	(0.099)	(0.169)	(0.068)
agedep	-0.546***	-0.708**	-0.667**	-0.461	0.271
	(0.188)	(0.351)	(0.317)	(0.352)	(0.188)
shurb	0.240**	0.120	0.133	0.821*	0.656*
	(0.107)	(0.180)	(0.365)	(0.461)	(0.387)
exp	-0.136**	-0.055	-0.057	0.184	0.005
	(0.054)	(0.061)	(0.067)	(0.134)	(0.144)
R-squared	0.349	0.358	0.165		
Wald test	<2.2e-16	<2.2e-16	<2.2e-7	0.14	<2.2e-16
Sargan test				0.124	0.001

Notes: Robust standard errors in parenthesis

*** p-value $\leq$ 0.01, ** p-value $\leq$ 0.05, * p-value $\leq$ 0.1

Table 3.8: Capital regressions: Middle income countries

Middle income countries N=280, dep. var.: investment share (ci), static					
	OLS	RE	FE	GMM	SYS-GMM
	(1)	(2)	(3)	(4)	(5)
enuse	0.157***	0.182***	0.069	0.682***	0.815***
	(0.030)	(0.065)	(0.076)	(0.231)	(0.161)
pop	-0.001	-0.039	-0.245*	-0.336	0.004
	(0.017)	(0.035)	(0.145)	(0.214)	(0.053)
popg	-1.583	2.291*	2.972**	1.133	1.959
	(1.346)	(1.187)	(1.233)	(1.345)	(2.689)
cg	-0.001	0.001	0.001	0.001	-0.001
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
openc	0.077*	0.167***	0.255***	0.289***	0.027
	(0.043)	(0.051)	(0.056)	(0.075)	(0.084)
agedep	-0.337**	-0.030	-0.115	0.149	0.150
	(0.143)	(0.177)	(0.209)	(0.228)	(0.139)
shurb	0.151**	0.027	0.045	-0.118	-0.687**
	(0.060)	(0.111)	(0.169)	(0.276)	(0.279)
exp	0.071	-0.179***	-0.125*	-0.154*	-0.225
	(0.075)	(0.062)	(0.058)	(0.088)	(0.173)
R-squared	0.337	0.184	0.087		
Wald test	<2.2e-16	<2.2e-16	0.001	<2.2e-16	<2.2e-16
Sargan test				0.001	0.001

Notes: Robust standard errors in parenthesis

*** p-value $\leq$ 0.01, ** p-value $\leq$ 0.05, * p-value $\leq$ 0.1

Table 3.9: Capital regressions: Low income countries

Low income countries N=212, dep. var.: investment share (ci), static					
	OLS	RE	FE	GMM	SYS-GMM
	(1)	(2)	(3)	(4)	(5)
enuse	0.337***	0.083	-0.108	0.192	0.001
	(0.093)	(0.169)	(0.232)	(0.274)	(0.256)
pop	-0.023	-0.054	-0.443*	-0.585	-0.046
	(0.035)	(0.069)	(0.226)	(0.364)	(0.063)
popg	-0.628	-0.678	-1.951	-1.542	-1.119
	(2.866)	(2.793)	(2.973)	(3.114)	(4.766)
cg	0.002	0.002	0.004	0.006**	0.003
	(0.002)	(0.002)	(0.003)	(0.002)	(0.004)
openc	-0.091	0.336***	0.443***	0.416***	-0.066
	(0.095)	(0.090)	(0.093)	(0.096)	(0.145)
agedep	0.513**	0.337	0.034	-0.524	-0.423
	(0.249)	(0.305)	(0.458)	(0.450)	(0.370)
shurb	-0.124	-0.148	0.021	0.462	0.061
	(0.091)	(0.163)	(0.253)	(0.455)	(0.211)
exp	2.321***	0.708	0.882	-0.568	1.150**
	(0.338)	(0.531)	(0.747)	(1.047)	(0.528)
R-squared	0.260	0.183	0.176		
Wald test	<2.2e-12	<2.2e-6	<2.2e-7	<2.2e-7	<2.2e-16
Sargan test				0.153	0.319

Notes: Robust standard errors in parenthesis

*** p-value ≤ 0.01, ** p-value ≤ 0.05, * p-value ≤ 0.1

Table 3.10: Capital regressions (dynamic)

All countries			**High income**		**Middle income**		**Low income**	
N=732			N=240		N=280		N=212	
	GMM	SYS-GMM	GMM	SYS-GMM	GMM	SYS-GMM	GMM	SYS-GMM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ci(-1)	0.401***	0.711***	0.581***	0.857***	0.224*	0.669***	0.423**	0.690***
	(0.145)	(0.062)	(0.174)	(0.132)	(0.128)	(0.094)	(0.178)	(0.154)
enuse	-0.294	0.162**	-0.154	0.082	0.511**	0.088	0.018	-0.288
	(0.254)	(0.078)	(0.207)	(0.051)	(0.249)	(0.088)	(0.261)	(0.250)
pop	-0.163	0.008	0.324	0.002	-0.615**	0.004	0.127	-0.009
	(0.201)	(0.012)	(0.250)	(0.009)	(0.271)	(0.018)	(0.345)	(0.034)
popg	0.459	-0.327	4.774***	-1.726	1.758	-0.892	-2.256	-6.408**
	(1.342)	(1.063)	(1.592)	(1.302)	(2.472)	(1.871)	(2.490)	(3.199)
cg	-0.001	-0.001	-0.001	-0.001	-0.002	-0.001	0.004**	0.000
	(0.002)	(0.001)	(0.007)	(0.001)	(0.004)	(0.001)	(0.002)	(0.001)
openc	0.332***	0.035	0.156	0.020	0.176*	0.048	0.209*	0.058
	(0.073)	(0.026)	(0.130)	(0.030)	(0.099)	(0.046)	(0.112)	(0.094)
agedep	-0.284	-0.021	-0.158	0.058	-0.051	-0.060	-0.510	0.074
	(0.208)	(0.057)	(0.178)	(0.125)	(0.308)	(0.100)	(0.333)	(0.210)
shurb	0.063	-0.148	-0.259	-0.077	0.080	-0.176*	0.273	0.062
	(0.219)	(0.098)	(0.206)	(0.120)	(0.271)	(0.100)	(0.270)	(0.121)
exp	0.080	0.041	0.024	-0.038	0.019	0.273**	-0.256	0.514*
	(0.124)	(0.065)	(0.094)	(0.060)	(0.146)	(0.122)	(0.562)	(0.311)
Wald test	<2.2e-8	<2.2e-16	<2.2e-16	<2.2e-16	0.001	<2.2e-16	<2.2e-16	<2.2e-16
Sargan test	0.002	0.002	0.250	0.132	0.166	0.186	0.602	0.535

Notes: Robust standard errors in parenthesis

*** p-value $\leq$ 0.01, ** p-value $\leq$ 0.05, * p-value $\leq$ 0.1

Table 3.11: Tests

Growth Regressions				
	Hausmann		Sargan difference	
	(1)	(2)*	(3)	(4)*
All countries	5.545e-07	1.800e-06	0.125	0.160
High income countries	0.001	0.001	0.085	0.123
Middle income countries	3.053e-05	1.243e-06	0.801	0.562
Low income countries	0.005	0.006	0.061	0.184

Capital regressions	static	static	dynamic
	Hausmann	Sargan difference	Sargan difference
	(5)	(6)	(7)
All countries	6.736e-14	0.274	0.186
High income countries	7.568e-08	0.649	0.107
Middle income countries	0.569	0.046	0.485
Low income countries	5.935e-15	0.660	0.649

Notes: All numbers reported are p-values

* Including human capital for growth regressions

Chapter 4

Stock Performance and Economic Growth: The Japanese Case*

Operating under unique macroeconomic conditions, the Japanese financial sector has attracted academic attention due to weak evidence for the momentum effect on the stock market. This paper relates the standard factor pricing models to growth expectations by testing for structural instability and by linking the profitability of the standard return-based risk factors (HML, SMB and WML) to economic growth. We find that the HML- and the WML-factor are statistically associated with economic growth. Accordingly, the description of stock returns by the usual risk factors is improved considerably when the estimations are conducted for subsamples representing different growth regimes, which particularly applies to the momentum strategy. The Japanese case illustrates the necessity of considering structural instability in relation to growth expectations, which is especially important for countries and in time periods with a sluggish economy.

*This chapter represents joint work together with Lucas Bretschger.

4.1 Introduction

This paper empirically investigates the performance of the standard factor pricing models against the background of economic growth in the Japanese economy. The interrelation of stock returns and macroeconomic factors in general is a research topic of highest priority. The need to understand the according economic mechanisms has increased with the latest global financial crisis as it is shown by current literature (Bouakez et al., 2013).

According to recent research in the field of empirical risk factor modeling, the book-to-market factor (HML) and the size factor (SMB) are associated with macroeconomic fundamentals, in particular with changes in economic growth expectations. Liew and Vassalou (2000) and Vassalou (2003) demonstrate that HML and SMB contain information about changes in growth expectations. In a recent contribution, Aretz et al. (2010) show that most macroeconomic factors are actually priced.

It is undisputed that the Fama-French three-factor approach (Fama and French 1993) and the momentum effect as proposed by Carhart (1997) had a widespread empirical success. Interestingly, Liu and Lee (2001) find that the momentum strategies in the Japanese stock market are not effective, where recent macroeconomic conditions differ quite substantially from other countries. In fact, the Japanese economy has experienced a period of absent or very low growth since the 1990s. Moreover, it is characterized by high government debt, amounting to 220 percent of GDP. At the same time, we observe a non-growing labor force, rising unemployment, decreasing savings rates, and near-to-zero (nominal deposit) interest rates. The Japanese experience can be interpreted as a prominent example for lasting stagnancy, which other leading economies might be confronted with in the future. Still, Japan is ranked among the world's largest economies in terms of real GDP and real GDP per capita; its stock market is the second largest in the world. Given the general finding that return-based risk factors contain information about macroeonomic risk it appears to be rewarding to re-evaluate the nature of the classical empirical risk factor models against the background of the unique macroeconomic conditions in Japan.

This paper contributes to the literature by relating the classical empirical factor pricing models for the Japanese stock market to economic growth. It is known that, before the non-growth period, Japan experienced a long phase of rapid growth and an impressive catch-up to the leading economies of the time. The switch to a new growth regime and the associated changes in asset pricing appear especially rewarding for study. Thus, the Japanese case allows us to learn more about the interrelation between stock performance and macroeconomic fundamentals. We assess the according sensitivity of the standard empirical asset pricing models in three ways: first, we test for the emergence of structural breaks during the full sample period from 1980 to 2009. Second, we directly relate the asset pricing models to macroeconomic fundamentals by regressing future economic growth on the return-based risk factors HML, SMB, and WML. Last, in line with the evidence for structural instability, we form subsamples, re-estimate the models, and compare the obtained results with the basic evidence from the full sample estimations.

We use newly constructed monthly data and risk factors (see Schmidt et al., 2011) for the time period 1984-2009. We compare the different base models, in particular the classical CAPM, the Fama-French three-factor model as well as the Carhart four-factor model including WML. We perform extensive robustness tests with regard to portfolio formation by altering the dimension from 5x5 to 4x4 and by alternating between equally weighted (EW) and value weighted (VW) portfolios.

In accordance with the Japanese macroeconomic development, we find that the factor pricing models exhibit structural instability and that the structural break occured in the late 1990s. The basic factor pricing regressions applied to the full sample period (1984-2009) reveal that the the null hypothesis of a zero intercept is rejected for the four-factor model which includes the momentum strategy. The re-estimations of the factor pricing models for two subsamples representing different growth regimes (the *growth period* 1984-1998 and the *stagnation period* 1998-2009) reveal that the null hypothesis of a zero intercept for the four-factor model cannot be rejected in either period. This indicates a considerable improvement of the return description with regard to the momentum strategy. Furthermore, we find evidence for the effectiveness of the momentum strategy in the *stagnation period* which stands in contrast to prior findings (see Liu and Lee, 2001). Further regression analysis shows that profitability of the HML and WML risk factor is linked to future economic growth. This appears to be revealing in two regards: first, prior literature (see Liew and Vassalou, 2000) only finds weak statistical association between risk factors and economic growth for Japanese data. Second, other than Liew and Vassalou (2000) and Griffin et al. (2003) we find that the momentum risk factor is linked to macroeconomic fundamentals. This confirms the finding that macroeconomic conditions cumulating in the structural break are crucial especially with regard to the momentum strategy. The results are found to be robust.

Evidence from prior studies suggests that especially the HML- and SMB-factor are statistically associated with macroeconomic fundamentals, e.g. see Aretz et al. (2007), Griffin et al. (2003), Hahn and Lee (2006), Kelly (2003), Lettau and Ludvigson (2001), Liew and Vassalou (2000), and Vassalou (2003). By considering cross-country differences Liew and Vassalou (2000) as well as Kelly (2004) seperately evaluate the Japanese case. Results are not unambigously clear. As mentioned above, Liew and Vassalou (2000), by evaluating the relation between return-based risk factors (HML, SMB, and WML) and future GDP growth, find only very weak statistical association. This result stands in contrast to most countries analysed in their study which generally reveals positive and significant parameters especially with regard to the HML and SMB factor. Kelly (2004) considers the relation between the two Fama-French-factors (HML and SMB) and future GDP growth as well as inflation. Other than Liew and Vassalou (2000), he finds that for Japan the SMB factor is positively and significantly associated with future economic growth and with inflation.

Past asset pricing studies for the Japanese stock market reflect a mixed performance of the standard risk models. Chan et al. (1991) conclude that the book-to-market ratio has a big impact on Japanese stock returns, which may also be due to to Japanese accounting standards; the cash flow yield and, to a minor extent, the size effect do also affect the stock performance. Kubota and Takehara (1996) reject the CAPM while Kubota and Takehara (1997) show that the Fama-French three-factor model captures the common risks in the Japanese stocks accurately. In contrast, Daniel et al. (2001) reject the Fama-French three-factor model but not the "characteristic" model, which links expected returns of assets to their characteristics which may have nothing to do with the covariance structure of returns. More recently, Long (2007) finds a reversal of the size effect for the period 1984-2004. Walid and Ahlem (2009) show that the CAPM is not an appropriate model for the Japanese market and Walid (2009) finds that both the firm size and book-to-market ratio are significantly related to average return premiums but suggests that there is stronger support to the characteristic model rather than the Fama-French three-factor model.

The remainder of the paper is organized as follows. Section 4.2 describes the data characteristics and the portfolio formation as well as the statistical framework in detail. In section 4.3, we elaborate on the interdependence of stock performance and the macroeconomic transition. We perform tests for structural changes and regression analysis including the typical risk factors and economic growth. Section 4.4 presents the empirical results for the factor pricing models applied to different samples representing different growth regimes. Section 4.5 summarizes and concludes.

4.2 Data and Methodology

4.2.1 Data Characteristics and Portfolio Formation

We use newly constructed market returns and risk factors based on Thomson Reuters Datastream and Thomson Reuters Worldscope data. A detailed documentation is given by Schmidt et al. (2011), who confirm the reliability of the thoroughly screened Thomson Reuters Datastream and Thomson Reuters Worldscope dataset by comparing the constructed market, value, size, and momentum risk factors with important benchmarks. We use monthly data from the Japanese stock market between 1984/7 and 2009/7. The number of firms was 403 in December 1984 while it amounted to 3558 in January 2009. The used risk factors are (for the details on the construction of the factors, see Schmidt et al. (2011), section 3.1):

- Fama-French risk factors: SMB (small minus big; related to the size, i.e. market capitalization), HML (high minus low; related to book-to-market value)
- Carhart's momentum factor: WML (winner minus loser)
- Market return: RM

In the case of Japan, the market return RM is highly and significantly correlated with the Tokio Stock Price Index (TOPIX); the estimated correlation is 0.996 in the VW (p-value $<$ 0.0001) and 0.838 in the EW (p-value $<$ 0.0001) case. We use the basic discount and loan rate (middle rate) as proxy for the risk-free rate Rf, the data are also from Thomson Reuters [1].

In order to analyze the returns, following the standard procedure in the literature, portfolios are formed each year with regard to size, book-to-market

[1] The usual proxy for the risk-free rate Rf is the Gensaki times series. However, we use the basic disount and loan rate due to better availability. The basic discount and loan rate is highly positively correlated with the Gensaki rate; the correlation estimation yields 0.978 (p-value $<$ 0.0001).

value and momentum. The breakpoints are set at the median values of the sorting variables, see also Schmidt et al. (2011), section 3.2. We consider the following portfolio-structures:

- 5x5 portfolios
  - 5 size-ranges (small to big) / 5 B/M-ranges (low to high)
  - 5 size-ranges (small to big) / 5 momentum-ranges (loser to winner)
- 4x4 portfolios
  - 4 size-ranges (small to big) / 4 B/M-ranges (low to high)
  - 4 size-ranges (small to big) / 4 momentum-ranges (loser to winner)

4.2.2 Statistical Framework

In the empirical analysis we consider the three standard versions of a factor pricing model. The corresponding time series representations of the asset pricing models are given by equations (4.1) to (4.3).[2] The dependent variable throughout the corresponding regressions is the excess return of portfolio i $(R_{it}-Rf_t)$ which is regressed on different combinations of the four risk factors described above. b_i, s_i, h_i and m_i are the accordant factor sensitivities for each portfolio i which are estimated from the time series regressions. N is the number of portfolios with the index i and T is the number of observations over time indexed by t. Thus, for $i = 1, 2, ..., N$, the following model specifications are considered:

$$R_{it} - Rf_t = a_i + b_i(RM_t - Rf_t) + e_{it}, \tag{4.1}$$

$$R_{it} - Rf_t = a_i + b_i(RM_t - Rf_t) + s_i SMB_t + h_i HML_t + e_{it}, \tag{4.2}$$

[2] In the remainder of this chapter we will refer to these three equations as model (1), model (2) and model (3).

$$R_{it} - Rf_t = a_i + b_i(RM_t - Rf_t) + s_i SMB_t + h_i HML_t + m_i WML_t + e_{it}. \quad (4.3)$$

a_i and e_{it} are asset return intercepts and disturbances, respectively. Model (1) is the Sharpe-Lintner version of the CAPM where excess portfolio returns are regressed on a constant and the excess market return only. Model (2) is often referred to as the Fama-French three-factor-model including additionally the HML and SMB factor. We refer to model (3) as the four-factor Carhart-like model where the momentum factor (WML) is added to describe portfolio returns. Model (1) is applied to size-B/M-sorted portfolios only, whereas model (2) and (3) are applied to size-B/M-sorted as well as to size-momentum-sorted portfolios[3].

In a first step, we descriptively analyze the sample moments of the variables involved. Then, by estimating the coefficients from the models above, we study common variation in portfolio returns. Additionally, we comparatively evaluate the precision of the different asset pricing specifications by the implication that each element of $\mathbf{a}=(a_1, a_2, ..., a_N)'$ is zero for a single model, which should be the case if the factors involved completely explain excess returns. Therefore, we will form a Wald test statistic of the null hypothesis $\mathbf{a}=\mathbf{0}$ against the alternative hypothesis $\mathbf{a}\neq\mathbf{0}$. That is, we test the joint hypothesis that all intercepts are zero. A rejection of the null hypothesis indicates a deviation from the exact factor pricing model.

Based on MacKinlay and Richardson (1991), inference is refined by applying a GMM approach. For every model (1) to (3), we jointly identify the parameters of interest by estimating a system of equation including all portfolios. Therewith, compared to single equation OLS, we are able to relax the assumptions that returns conditional on the factor realizations are IID through

[3] When not mentioned explicitly, the results are reported for model (2) applied to size-B/M-sorted portfolios and model (3) applied to size-momentum-sorted portfolios.

time.
The analysis is focused on the 4x4 sorted portfolios for value weighted returns. We will check the result's robustness by additionally applying the described framework on equally weighted returns as well as on 5x5 sorted portfolios.

Before analysing the factor pricing models in detail, following Bai and Perron (2003) and Andrews and Ploberg (1994), we test for structural instability of the models and construct subsamples for the empirical identification of risk factors. In line with Liew and Vassalou (2000) we apply further regression analysis in order to investigate the interdependence between the profitability of the factor-mimicking portolios and macroeconomic measures. More detailed information is given in section 4.3.

4.3 Macroeconomic Fundamental

Japan experienced a remarkable transition in macroeconomic fundamentals during the 1990s. For three decades, starting from 1960, aggregate production grew rapidly and Japan established itself as the world's second largest economy which is often referred to as the post-war economic miracle. During the 1980s low interest rates, high stock and real asset prices led to a crash of the Tokyo Stock Exchange in the early 1990s. Growth slowed considerably during the late 1990s initiating a persistent period of economic sluggishness characterized by absent or very low growth, high government debt, a non-growing labor force, rising unemployment, and near-to-zero (nominal deposit) interest rates. Accordingly, it has been broadly argued that Japan fell into a liquidity trap during this transitional period, a situation with low interest and high saving rates, rendering monetary policy ineffective (see Krugman, 1998).
The interrelation between macroeconomic factors and stock returns has been empirically investigated for a long time, e.g. see Chan et al. (1985) and Chen et al. (1986). Generally, it is found that investment opportunities

and therewith stock pricing properties are closely related to the macroeconomic environment. The development of economic fundamentals in Japan as described above therefore suggests that the structural properties of the empirical asset-pricing models do not remain constant throughout the full sample period from 1980 to 2009. We investigate this issue thoroughly in the following section.

4.3.1 Structural Change

In this section, we statistically address the problem of possible structural changes with regard to the three factor pricing models (1), (2), and (3). In line with Chow (1960), we compute a test-statistic for every conceivable single breaking point in order to test whether the coefficients of two resulting factor pricing models are different. As in chapter 2, following Andrews and Ploberg (1994), we reject the null hypothesis of structural stability if the supremum of these statistics exceeds a certain critical value. The test statistic follows an F-distribution. The test is applied for every model (1), (2), and (3) and for every portfolio $i = 1, 2, ..., N$ based on OLS estimations. From table 4.1 we derive that, apart from two exceptions, the null hypothesis of no structural change is rejected in any case. We determine the exact number of breaks by choosing the models with the minimal Bayesian information criterion (BIC) throughout different number of breakpoints. For 50% of the analyzed models the BIC is minimal at a single breaking point.

In the next step, given the evidence for one single breaking point, we assess the timing of the structural break by using the dating procedure described by Bai and Perron (1998, 2003). Consider the factor pricing regression (4.3) for portfolio i with l being the number of breakpoints:

$$R_{it} - Rf_t = a_{ij} + b_{ij}(RM_t - Rf_t) + s_{ij}SMB_t + h_{ij}HML_t + m_{ij}WML_t + e_{it}, \tag{4.4}$$

with $t = T_{j-1} + 1, ..., T_j$ for $j = 1, ..., l + 1$. For $l = 1$, T_1 denotes the unknown single breakpoint. Following Bai and Perron (2003), we estimate $\hat{T}_1$ based on the least-square principle such that $\hat{T}_1 = \underset{T_1}{\text{argmin}} S_T(T_1)$ with the minimization being taken over all possible sample partitions. S_T denotes the sum of squared residuals. Although the test is conducted for every regression ((1), (2), and (3)) and portfolio ($i = 1, 2, ..., N$) separately, we generalize the specific results from the 64 models such that we get a representative single breaking point. That is, in order to keep the design of the empirical analysis straightforward, we will eventually apply the identical sample segmentation periods to every model under examination.

The calculated sample segmentations are presented in table 4.2. In roughly 65% of the models the breaking point lies in the time period between 1997 and 2000. Excluding model (1) from the analysis this rate increases to 75% indicating a high degree of homogeneity in structural behavior throughout the models (2) and (3).

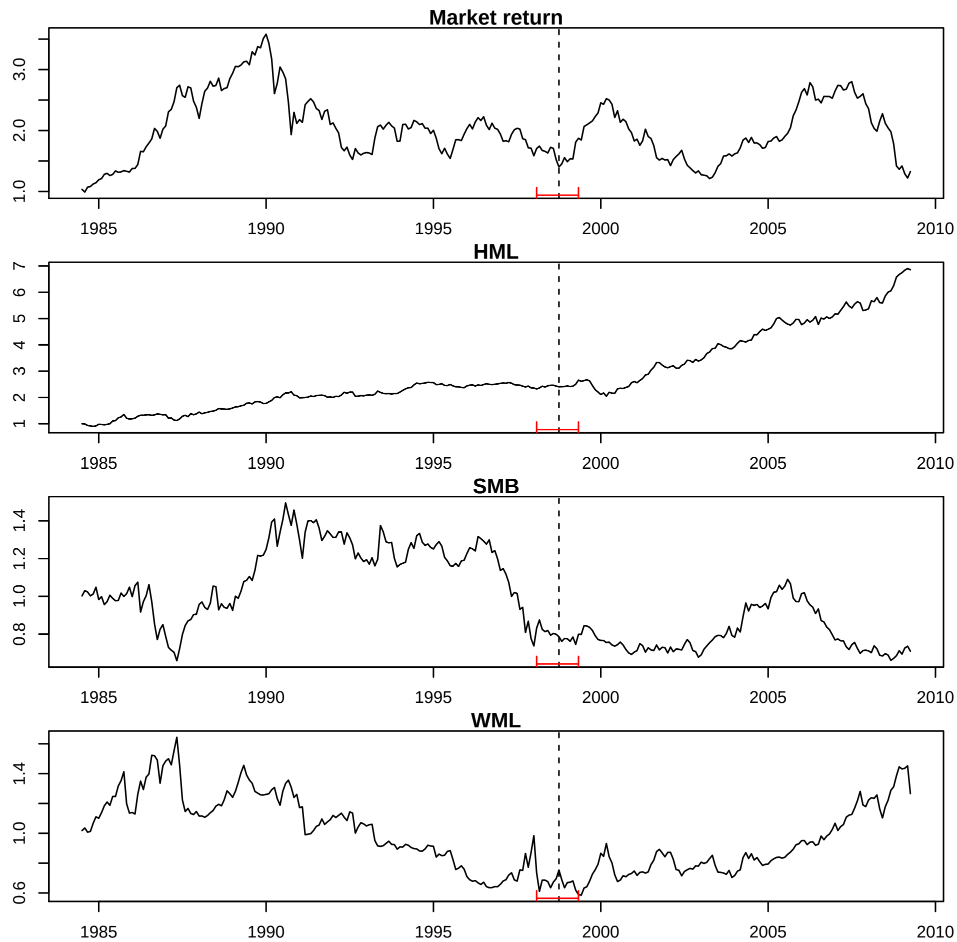
In line with the macroeconomic development in Japan, we therewith identify a structural break between the years 1997 and 2000. Whereas the terminal point of the post-war economic miracle is usually identified by the bursting of the asset price bubble in the early 1990s, the stock pricing process appears to require some time to adjust to the new macroeconomic regime. In fact, the second half of the 1990s is characterized by solvency problems of some of Japan's major banks, such as the failure of the Long Term Credit Bank of Japan (LTCB) which is considered the largest default during the 1990s (Nakaso, 2001). The nationalization of the LTCB that followed in October 1998 publicly demonstrated the extent of the financial crisis which possibly affected investment behaviour in a crucial way.

Accordingly, fixing the breaking point at October 1998, we can visually recover the structural change by looking at the four cumulated risk factors in figure 4.1. The market return in the upper section shows only a slight structural modification, whereas the two segments are clearly distinguishable for the other three risk factors. For the HML factor, the process shows a rather stationary behavior in the first period, whereas the second period seems to be governed by a positive trend. Compared to the first period, the SMB factor fluctuates around a lower average in the second period. For the WML factor, we can assess a distinct break around the year 1998 with the factor changing from a negative to a positive trend. The visual assesment shows that the cumulative risk factors exhibit a delayed reaction to the bursting of the asset price bubble.

4.3.2 Return-Based Risk and the Macroeconomy

In this section, we provide empirical evidence for the interrelation between asset performance and the macroeconomy for the Japanese case. The occurrence of structural model instability during the transitional period in the Japanese economy found in the previous section indicates that the asset-pricing process represented by the CAPM, Fama-French- and Carhart-model is crucially influenced by macroeconomic fundamentals. Merton's (1973) intertemporal CAPM (ICAPM) provides a possible analytical framework for interpreting investors' decisions and asset returns in a macroeconomic context. Whereas in the static Sharpe-Lintner CAPM setting agents only care about the mean and variance of one-period portfolio returns, the dynamic specification of the ICAPM extends agents' decision-making process by including future investment and consumption opportunities. Consistent with expected utility maximization, agents will allocate consumption over time where additional risk is imposed by the uncertainty coming from changing investment and consumption opportunities. Thus, apart from the market betas, asset pricing representations require additional factors in order to adequately capture the risk related to future state variables. According to Kelly (2003), macroeconomic measures related to labor income or prices of consumption goods

Figure 4.1: Breakingpoints and cumulated risk factors



can be interpreted as state variables as they reflect the business climate and therefore approximate shifts in agents' marginal utility. Transferred to an empirical context, Fama and French (1993) argue that their return based risk factors (HML and SMB) reflect such state variables: the higher returns on small stocks and high book-to-market stocks mirror state variables that generate undiversified risk in returns that is priced seperately from the risk related to market returns (Fama, French; 2004). Given the presumption that macroeconomic risk is contained in the factor-mimicking portfolios, we expect the profitability of HML, SMB and WML to be correlated with macroeconomic measures.

Based on the finding that the investigated asset-pricing models exhibit structural breaks within the period 1997 to 2000, we loosly inferred that this parameter instability is driven by the drastic change in the Japanese macroeconomic fundamentals during the 1990s. Accordingly, from figure 4.1, we observe that the risk factors which serve as regressors in the asset pricing models are subject to a noticeable alteration during this transitional period. In line with the descriptions above, the performance of the return-based risk factors may reflect changes in fundamental risk rising from changing investment and consumption opportunities. A long-term shift in macroeconomic fundamentals and a corresponding sustained change in investment behaviour such as in the Japanese example therewith represent a possible challenge to the assumption of parameter stability in the basic factor pricing models calling for more flexible model specifications.

Interpreting growth expectations as a key indicator for macroeconomic risk, following Kelly (2003) and Liew and Vassalou (2000), we test whether the profitability of HML, SMB and WML is associated with future economic growth. Using multivariate regression we re-evaluate the existing findings for the Japanese case against the background of structural breaks and with newly constructed factor-mimicking portfolios. The regressions use quarterly data from the third quarter in 1984 to the second quarter in 2009 and are of the form

$$G_{(t,t+4)} = a + b * RM_{(t-4,t)} + c * Factor_{(t-4,t)} + e_t, \tag{4.5}$$

$$G_{(t,t+4)} = a + b * RM_{(t-4,t)} + c * SMB_{(t-4,t)} + d * HML_{(t-4,t)} + f * WML_{(t-4,t)} + e_t. \tag{4.6}$$

Next year's growth in production $G_{(t,t+4)}$ is regressed on past year's annual portfolio return where $Factor_{(t-4,t)}$ in (4.5) represents HML, SMB, or WML respectively. Portfolio returns are the same as before and GDP growth is approximated by the index of industrial production as provided in the MEI database of the OECD. The equations are estimated by OLS where we use Newey-West standard errors to correct for serial correlation.

Looking at correlation coefficients in table 4.3 we see that the SMB factor as well as the market return are significantly and positively related to future economic growth, which is consistent with Kelly's (2003) finding. However, using regression (4) for each factor seperately and controlling for the market return we find that in addition to SMB which is still positive and significant the HML factor is also positively (significant at a 10% level) and WML is negatively but not significantly related to future economic growth. The coefficient for market return remains positive and significant. Using regression specification (5) we find that the market return and the HML factor are significant and positive. WML is negatively (significant at a 10% level) associated with future economic growth. Prior studies such as Liew and Vassalou (2000) and Griffin et al. (2003) find that the momentum factor is unable to explain variation in macroeconomic fundamentals.

The interrelation of the HML factor and future economic growth can be understood on the grounds of firm-level characteristics. Following Kelly (2003), changes in real economic growth are associated with changes in systematic firm-level risk related to size and book-to-market ratio. The negative association of the WML-factor profitability and future economic growth can be interpreted in line with the basic momentum literature. A possible reason for the existence of a momentum effect is that the market underreacts to information (see e.g. Jagadeesh and Titman). The negative relation to future GDP growth indicates that this underreaction is stronger, the lower are expectations regarding the future economic state.

We conclude that for the Japanese case return-based risk factors, especially

SMB and WML, are linked to future economic growth indicating that factor-mimicking portfolios reflect macroeconomic risk. Relating this finding to structural breaks, we deduce that the drastic and persistent change in macroeconomic fundamentals might have altered the structural properties of the asset-pricing processes through the changed dynamics in the profitibality of return-based risk factors. In the next section, we therefore apply the three empirical factor pricing models to the full as well as to a subdivided data set. Two subsamples are constructed representing two distinct macroeconomic regimes: the *growth period* from 1984/7 to 1998/1 and the *stagnation period* from 1998/2 to 2009/7. We fix the breaking point at January 1998 which is consistent with the evidence in the last section.

4.4 Empirical Evidence: Factor Pricing Models

4.4.1 Basic Estimations: Full Period (1984 -2009)

The calculated means for the portfolios formed on size and book-to-market equity in table 4.4 show that the returns increase monotonically and consistently from the lowest to the highest portfolio. Accordingly, as it is shown in table 4.5 the difference between the return of the highest B/M minus the lowest B/M portfolio is significantly different from zero, which indicates a positive relation between average return and B/M equity. We also find some evidence that there is a negative relationship between returns and size, but in a less consistent manner. Specifically, the returns in the biggest portfolio seem to be greater than the next smaller portfolio return. Consequently, for each category, the difference between the smallest and biggest portfolio is not statistically different from zero.

Looking at the portfolios formed on size and momentum in the lower sections of table 4.4 and table 4.5, we observe that there is no clear evidence for a momentum effect from the average means. The difference between the returns of the winner and the loser portfolio is not statistically different from zero through every size category. However, forming the portfolios on a size-momentum-basis instead of the size-B/M basis reveals a clear negative relationship between average returns and size, with the difference of the return from the smallest minus the biggest portfolio being partly significantly different from zero.

Generally, the descriptive statistics correspond with established findings from the Japanese stock market, confirming the reliability of the newly constructed market returns and risk factors.

The estimated parameters of model (1) are visualized in table 4.6. The b_is have the expected sign and magnitude, ranging around 1. They are highly significant for every portfolio i. However, the b_is cannot sufficiently explain the differences in returns between the portfolios. Moreover, in 4 out of 16 cases, the null-hypothesis of a_i being equal to zero is rejected. The hypothesis for a_i being jointly zero throughout all portfolios is rejected indicating a misspecification of the CAPM.

As expected, we see from the left section of table 4.7 that the three factors in model (2) capture common variation in stock returns. The b_is are all highly significant. s_i and h_i (except for a few exceptions) are also significant. The slopes of HML_t and SMB_t are related to size and B/M respectively. s_i decreases with size and h_i increases with a higher B/M ratio explaining the variation in portfolio returns described in the descriptive analysis. All estimates for a_i are significantly different from zero, and the joint hypothesis cannot be rejected which, compared to model (1), indicates an improvement of the asset pricing specification.

The 16 estimated b_is in model (1) range from 0.773 and 1.077 with a sample variance of 0.006, whereas in model (2) they lie between 0.907 and 1.090 with a sample variance of 0.002. This shows that with regard to the excess market return the factor sensitivities in the three-factor model exhibit some form of convergence over the different portfolios. According to Fama and French (1993) "(...) Adding SMB and HML to the regressions collapses the betas for stocks toward 1 (...). This behavior is, of course, due to the correlation between the market and SMB or HML." Consistently, we see in the correlation matrix from table 4.4 that the correlation between HML and SMB is low and not significant, whereas it is significant between RM and HML. Furthermore, the estimates for b_i seem to be systematically lower in model (1). Adding the momentum factor to the regression (2) does not change the results. b_i, s_i and h_i are robust with regard to sign and magnitude. According to table 4.12, the values of the Wald statistics of the null hypothesis are slightly higher for the four-factor model, indicating some improvement in the description of portfolio returns. The momentum factor WML is negative and partly significant, especially for the portfolios with a low B/M ratio[4].

The results for model (3) are presented in the right hand section of table 4.7. The portfolios are now formed on size and momentum factors. The b_is are still highly significant and range around 1. The s_i and h_i coefficients are also significant in most cases. Both factor sensitivities, however, show less significance for the biggest portfolio category.

Except for two portfolios, the momentum-factor m_i is significant in explaining common variation. We observe that m_i is monotonically increasing from loser to winner portfolios, which, however, is not reflected in the returns described in the descriptive statistics above. From the regression results, we find that for the winner portfolio the h_is are consistently lowest within the corresponding size category.

Remarkably, the null hypothesis **a=0** is rejected at a very low significance level. Thus, although model (3) consists of four risk factors, the asset pricing process is described poorly (or in an incomplete manner, to put it more mildly).

[4]The corresponding results can be obtained from the authors upon request.

4.4.2 Sub-Sample Estimations

Considering the explanatory returns, from the descriptive statistics presented in table 4.8 and table 4.9, we see that the calculated moments are quite similar to the full period sample. With regard to the dependent returns, we see that for the portfolios formed on size and book-to-market equity returns increase from the lowest to the highest portfolio for both time periods.

The returns of the portfolios formed on size and momentum show some revealing properties. From table 4.10 and table 4.11 we find evidence for the momentum strategy in period 2 as the winner portfolio significantly outperforms the loser portfolio for two size-categories. On the other hand, there is evidence for a reversed momentum effect in period 1, with the loser portfolio exhibiting a higher average return than the winner portfolio for one size-category at the 15% significance level. Thus, based on this considerations, the breaking point seems to be crucial with regard to the momentum strategy. Contrary to Liu and Lee (2001), we are able to show that stock returns appear to follow a continuation pattern after 1998 which is reflected in the positive difference between the winner and the loser portfolio.

For model (1) presented in table 4.13 we cannot reject the null hypothesis $\mathbf{a}=\mathbf{0}$ for period 1, indicating a satisfying performance of the CAPM between 1984 and 1998. As for the full period, the same hypothesis is rejected for period 2. The estimates for b_i are systematically higher in period 2.

Similar to the estimates for the full time period, the inclusion of the SMB and HML risk factors in table 4.14 improves the performance of the asset pricing model. In both cases we cannot reject the null hypothesis $\mathbf{a}=\mathbf{0}$.

In contrast to the full period estimates, the null hypothesis $\mathbf{a}=\mathbf{0}$ cannot be rejected for model (3) in both periods. That is, compared to the last section, we find a clear improvement of the four factor model when applied separately for the two time periods. We see from table 4.15 that the estimated factor sensitivities b_i, s_i, h_i and m_i behave in a similar way as for the full period. Similar to model (2), s_i and h_i tend to be lower for period 1. The momentum sensitivity m_i increases monotonically from loser to winner portfolios for both time periods. As for the full period, the application of the three-factor model to size-momentum-sorted portfolios does not notably change the results with regard to b_i, s_i and h_i[5]. Other than for the full period, the comparison of the test statistics in table 4.12 reveals a more considerable improvement in the description of portfolio returns when the momentum factor is included.

The regression results as well as the descriptive statistics are mostly robust when applied to equally weighted portfolios sorted on a 4x4 and 5x5 basis as well as for value weighted portfolios sorted on a 5x5 basis. [6]

4.4.3 Summary and Discussion

We investigated the effect of the macroeconomic transition in Japan on the standard empirical factor pricing models (CAPM, Fama-French, and Carhart). In line with observational evidence from macroeconomic indicators we statistically detect structural model instability during the second half of the 1990s. Further regression analysis reveals that profitability of the risk-factor portfolios, especially SMB and WML, are linked to the macroeconomic environment (represented by future economic growth). This emphasizes the sensitivity of the asset-pricing models to macroeconomic fundamentals.

It turns out that splitting the data set into a *growth period* from 1984/7 to 1998/10 and a *stagnation period* from 1998/11 to 2009/7 is especially important in several respects. First, other than for the full period, the hypothesis

[5]The corresponding results can be obtained from the authors upon request.

[6]An exception are the equally weighted portfolios, where the null hypothesis **a=0** for model (3) in period 2 is rejected, weakening somewhat the general result of the four factor model performing better when applied to the subdivided sample. Furthermore, for equally weighted 5x5 portfolios in period 2, we do not find significant evidence for the momentum strategy calculated from the difference in average portfolio returns. The corresponding tables containing the complete set of calculations and estimations can be obtained from the authors on request.

for the intercepts being jointly zero for the CAPM cannot be rejected for the first period indicating a more satisfying performance when applied to the subsamples. Considering the momentum strategy, descriptive statistics reveal that for the full period the difference between the returns of the winner and the loser portfolio is not statistically different from zero. Looking at the single subsamples, however, we find a reversed effect in the first period with the loser portfolio exhibiting a higher average return than the winner portfolio. Interestingly, for the second period we find evidence that the (Winner-Losers)-portfolio is positive for two out of four size groups. This result stands in contrast to Liu and Lee (2001) who suggest that Japanese stock prices tend to exhibit a reversal pattern. They empirically evaluate monthly stock returns for the period from 1975 to 1997. Their results are robust to the application to a subdivided dataset: they consider the bull-market (1975-1989) and the bear-market (1990-1997). On the contrary, we suggest that the stock pricing processes need more time to adjust to the changing macroeconomic environment. We show that after 1998 stock prices exhibit a continuation pattern consistent with momentum-evidence both from U.S. and European markets.

From the regression analysis we find that the Wald test for the intercepts being jointly zero cannot be rejected for the Carhart model after splitting the data sample, whereas the same test indicates that intercepts are not jointly zero for the full period. Hence, considering the structural break, the standard four factor model including the momentum factor explains returns more adequately. This finding is reflected in the significant association between the WML factor and future GDP growth.

Our results can be interpreted in line with a branch of literature examining the stability of betas in factor pricing models.[7] The general finding is that portfolio betas behave counter-cyclical, they increase (decrease) when the market is bearish (bullish). We contribute to the literature by showing

[7] See Levy (1974), and Fabozzi and Francis (1977) for seminal contributions.

that this instability in the sample coefficients is related to macroeconomic fundamentals. For the Japanese case, the momentum strategy appears to be especially sensitive in this regard.

4.5 Conclusions

Using a new set of data and risk factors, we analyze the stock performance in the Japanese market against the background of a transition in the growth regime. Specific testing reveals that we have a structural break in 1998, indicating the change from a growing to a mainly stagnant economy. This insight is reinforced by the finding that risk factors, especially HML and WML, are statistically associated with future economic growth. That is, a break in macroeconomic fundamentals may have altered the structural properties of the asset-pricing processes through the changed dynamics in return-based risk factors' profitability.

Applying the conventional factor pricing models to a subdivided data set representing two distinct growth-regimes reveals a considerable improvement of the empirical description of portfolio returns compared to a full period estimation. This finding especially applies to the Carhart four-factor model. Furthermore, we find evidence for stock prices exihibiting a continuation pattern after 1998 consistent with momentum-evidence both from U.S. and European markets.

We conclude that given the current sluggishness of the world economy, researchers and practitioners should be increasingly alert for structural breaks, following the growth expectations in the economy. Overall, the paper is additional proof for the robustness of the Fama-French approach, for periods of both high and low economic growth. It also shows that for the case of Japan the Carhart-model including the momentum factor is especially vulnerable when macroeconomic conditions change. This may represent a possible explanation for the absence of momentum-effects detected in prior studies on the Japanese stock market.

It would be interesting to see whether the new evidence for the Japanese market can be corroborated when performing similar tests for other markets. In addition, the international links between financial markets with regard to the momentum effect would be interesting to study. This is left for further research.

4.6 Appendix

Table 4.1: Structural stability: F tests

	Model (1)				**Model (2)**			
p-values	Low	2	3	High	Low	2	3	High
Small	≤2.2e-03	0.018	0.001	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03	*0.064*
2	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03
3	≤2.2e-03	0.030	0.024	0.050	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03
Big	≤2.2e-03	≤2.2e-03	0.038	*0.872*	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03

	Model (3)*				**Model (3)**			
p-values	Low	2	3	High	Loser	2	3	Winner
Small	≤2.2e-03	≤2.2e-03	≤2.2e-03	0.023	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03
2	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03
3	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03	≤2.2e-03
Big	≤2.2e-03	≤2.2e-03	≤2.2e-03)	≤2.2e-03	0.018	≤2.2e-03	≤2.2e-03	≤2.2e-03

Notes: p-value¿0.05 are set in italics

* Applied to portfolios formed on size and book-to-market equity.

Table 4.2: Breaking points

	Model (1)				**Model (2)**			
p-values	Low	2	3	High	Low	2	3	High
Small	1990(3)	1987(8	1990(3)	1990(3)	1997(8)	1990(10	1987(2)	1986(11)
2	1990(5)	1987(8)	1990(3)	1990(8)	1997(5)	1999(2)	1999(10)	1998(10)
3	1989(1)	1998(12)	1999(10)	1990(8)	1997(10)	1997(6)	1999(10)	1997(2)
Big	1990(8)	1986(11)	2000(3)	1999(6)	1998(8)	1993(4)	2000(5)	2002(10)

	Model (3)*				**Model (3)**			
p-values	Low	2	3	High	Loser	2	3	Winner
Small	2002(10)	1990(4	1990(10)	1999(9)	1992(11)	1998(2)	1998(7)	1999(10
2	2000(12)	1999(2)	1999(10)	1999(4)	1999(11)	1998(3)	1998(3)	1999(10
3	1999(5)	1997(6)	1999(10)	1997(2)	1999(10)	1998(7)	1997(7)	1997(10)
Big	1999(1)	1993(5)	2000(6)	2002(10)	2000(6)	1987(2)	2000(6)	1997(5)

Notes: The breakingpoint is defined as the last observation of the first period.

* Applied to portfolios formed on size and book-to-market equity.

Table 4.3: GDP growth and return-based risk factors

Correlations

	RM	SMB	HML	WML
GDP growth	0.423**	0.283**	-0.123	-0.128

Regression results

	Slope coefficients				t-values			
	RM	SMB	HML	WML	RM	SMB	HML	WML
	0.089	0.114			2.031	2.193		
	0.110		0.181		2.675		2.129	
	0.101			-0.084	2.262			-1.140
	0.123	-0.000	0.208	-0.110	3.162	-0.007	2.054	-1.657

Notes:

For correlation matrix: ** p-value≤0.01, * p-value≤0.05

Table 4.4: Descriptive statistics: 16 portfolios, full period (1984-2009)

Explanatory returns

	Sample moments				Correlation matrix				
	Rm	SMB	HML	WML		RM	SMB	HML	WML
Mean	0.115	-0.027	0.690	0.206	RM	1.000	-0.088	-0.244*	-0.177*
Median	0.280	-0.173	0.636	0.766	SMB	-0.088	1.000	0.088	-0.236*
Maximum	18.405	15.014	10.522	15.058	HML	-0.244*	0.088	1.000	0.045
Minimum	-22.000	-14.711	-10.777	-25.299	WML	-0.177*	0.236*	0.045	1.000
t-value	0.344	-0.110	4.093	0.718					

Dependent returns: portfolios formed on size and book-to-market equity

	Mean				t values			
	Low	2	3	High	Low	2	3	High
Small	0.139	0.421	0.526	0.652	0.294	1.023	1.356	1.697
2	-0.201	0.806	0.403	0.582	-0.460	0.208	1.050	1.474
3	-0.128	0.221	0.366	0.517	-0.310	0.598	1.022	1.330
Big	-0.203	0.404	0.646	0.593	-0.542	1.177	1.884	1.648

Dependent returns: portfolios formed on size and momentum

	Mean				t values			
	Losers	2	3	Winner	Losers	2	3	Winner
Small	0.688	0.704	0.886	0.605	1.449	1.746	2.331	1.554
2	0.155	0.363	0.477	0.347	0.336	0.924	1.296	0.913
3	0.047	0.183	0.372	0.352	0.101	0.473	1.050	0.954
Big	0.044	0.174	-0.118	0.345	0.091	0.453	-0.342	0.929

Notes:

For correlation matrix: ** p-value≤0.01, * p-value≤0.05

Table 4.5: Differences in means for extreme portfolios: full period (1984-2009)

Portfolios formed on size and book-to-market equity

	Difference in means				t values			
	Small	2	3	Big	Small	2	3	Big
High-low	0.514	0.784	0.645	0.796	2.374	3.561	3.473	2.878
	Low	2	3	High	Low	2	3	High
Small-big	0.342	0.017	-0.012	0.059	0.908	0.053	-0.430	0.191

Portfolios formed on size and momentum

	Difference in means				t values			
	Small	2	3	Big	Small	2	3	Big
Winner-losers	-0.082	0.192	0.305	0.302	-0.290	0.714	1.0281	0.737
	Loser	2	3	Winner	Loser	2	3	Winner
Small-big	0.644	0.531	1.004	0.260	1.953	1.688	3.435	0.748

Table 4.6: Model (1): 16 portfolios, full period (1984-2009)

	b_i				$t(b_i)$			
	Low	2	3	High	Low	2	3	High
Small	0.977	0.850	0.819	0.773	13.037	12.687	12.961	12.362
2	0.955	0.912	0.893	0.894	14.319	15.936	14.096	15.526
3	1.034	0.953	0.904	0.979	19.088	17.552	21.468	20.099
Big	1.077	0.961	0.960	0.870	37.069	27.090	33.160	18.749

	a_i				$t(a_i)$			
	Low	2	3	High	Low	2	3	High
Small	0.000	0.003	0.004	0.006	0.081	0.979	1.404	1.641
2	-0.003	0.000	0.003	0.005	-1.058	-0.096	1.189	1.609
3	-0.002	0.001	0.003	0.004	-1.175	0.582	1.144	1.562
Big	-0.003	0.003	0.005	0.005	-2.642	2.009	3.965	2.323

Linear hypothesis test for $\mathbf{a}=\mathbf{0}$

p-value	0.001

Table 4.7: Model (2) and (3): 16 portfolios, full period (1984-2009)

Model (2)									**Model (3)**								
	b_i				$t(b_i)$					b_i				$t(b_i)$			
	Low	2	3	High	Low	2	3	High		Loser	2	3	Winner	Loser	2	3	Winner
Small	1.057	0.937	0.911	0.907	33.849	33.307	31.853	34.244	Small	0.977	0.905	0.917	0.915	35.390	27.052	22.601	23.624
2	0.990	0.989	0.985	1.037	35.254	46.340	44.694	69.298	2	1.011	0.806	0.965	0.966	46.478	38.504	30.485	27.895
3	1.059	1.013	0.987	1.090	40.815	31.579	39.312	41.308	3	1.063	0.983	0.978	1.025	38.914	27.103	26.641	25.749
Big	1.016	0.968	1.005	0.973	59.105	30.218	43.264	30.036	Big	1.010	0.951	0.964	1.058	29.347	26.515	28.046	42.885

	s_i				$t(s_i)$					s_i				$t(s_i)$			
	Low	2	3	High	Low	2	3	High		Loser	2	3	Winner	Loser	2	3	Winner
Small	1.235	1.069	0.977	0.962	24.051	36.578	21.759	25.511	Small	1.097	1.028	0.986	1.012	27.721	24.925	23.192	19.540
2	1.105	0.891	0.891	0.879	22.321	33.253	42.908	41.878	2	0.886	0.808	0.808	0.919	30.060	27.658	24.922	22.012
3	0.744	0.591	0.548	0.589	15.996	17.547	15.421	16.263	3	0.611	0.563	0.539	0.610	14.910	14.798	14.248	12.725
Big	-0.138	-0.201	-0.011	-0.097	-5.787	-4.459	-0.302	-1.864	Big	-0.050	-0.138	-0.084	-0.066	-0.954	-2.331	-1.877	-1.351

	h_i				$t(h_i)$					h_i				$t(h_i)$			
	Low	2	3	High	Low	2	3	High		Loser	2	3	Winner	Loser	2	3	Winner
Small	0.001	0.152	0.240	0.582	0.013	1.756	2.709	7.230	Small	0.219	0.261	0.291	0.151	3.263	4.106	3.819	1.682
2	-0.292	0.161	0.283	0.702	-3.924	3.495	5.405	11.219	2	0.227	0.306	0.306	0.209	4.298	4.969	4.749	2.637
3	-0.186	0.180	0.391	0.597	-2.794	2.7898	5.391	8.222	3	0.257	0.361	0.321	0.149	3.338	6.206	4.592	1.870
Big	-0.419	0.164	0.372	0.892	-6.636	2.387	6.633	9.198	Big	0.060	0.070	0.325	0.052	0.706	0.681	4.101	0.941

										m_i				$t(m_i)$			
										Loser	2	3	Winner	Loser	2	3	Winner
									Small	-0.484	-0.173	0.097	0.285	-11.628	-4.726	1.929	4.444
									2	-0.551	-0.236	0.032	0.267	-16.829	-6.544	0.748	4.549
									3	-0.659	-0.303	-0.014	0.230	-15.997	-7.685	-0.280	5.117
									Big	-0.914	-0.420	-0.125	0.453	-15.676	-8.658	-2.179	8.613

	a_i				$t(a_i)$					a_i				$t(a_i)$			
	Low	2	3	High	Low	2	3	High		Loser	2	3	Winner	Loser	2	3	Winner
Small	0.001	0.002	0.003	0.002	0.247	1.361	1.604	1.205	Small	0.006	0.005	0.006	0.004	3.947	3.300	3.278	1.724
2	-0.001	-0.001	0.001	0.000	-0.491	-0.950	0.953	0.026	2	0.000	0.001	0.002	0.001	0.162	0.851	1.092	0.342
3	-0.001	-0.000	-0.000	-0.000	-0.614	-0.025	-0.017	-0.030	3	-0.001	-0.001	0.001	0.001	-0.645	-0.733	0.333	0.538
Big	-0.000	0.002	0.003	-0.001	-0.321	1.241	2.076	-0.758	Big	0.001	0.001	-0.004	0.001	0.424	0.522	-2.765	0.660

Linear hypothesis test for **a=0**

p-value	0.279	p-value	¡2.2e-03

Table 4.8: Descriptive statistics: 16 portfolios, period 1 (1984-1998)

Explanatory returns

	Sample moments					Correlation matrix			
	Rm	SMB	HML	WML		RM	SMB	HML	WML
Mean	0.220	-0.072	0.574	0.107	RM	1.000	-0.088	-0.221*	-0.099*
Median	0.253	0.194	0.510	0.345	SMB	-0.088	1.000	0.119	-0.344*
Maximum	18.405	15.014	10.522	15.058	HML	-0.221*	0.119	1.000	0.088
Minimum	-22.000	-14.711	-10.777	-25.299	WML	-0.099*	-0.344*	0.088	1.000
t-value	0.468	-0.192	2.407	0.284					

Dependent returns: portfolios formed on size and book-to-market equity

	Mean				t values			
	Low	2	3	High	Low	2	3	High
Small	0.191	0.412	0.524	0.725	0.295	0.678	0.906	1.265
2	-0.161	0.005	0.306	0.507	-0.269	0.010	0.531	0.892
3	-0.316	0.131	0.221	0.314	-0.560	0.240	0.418	0.569
Big	-0.107	0.529	0.493	0.593	-0.193	1.013	1.017	1.213

Dependent returns: portfolios formed on size and momentum

	Mean				t values			
	Loser	2	3	Winner	Loser	2	3	Winner
Small	0.673	0.765	0.988	0.273	1.014	1.231	1.685	0.482
2	0.165	0.374	0.408	0.088	0.261	0.634	0.727	0.1661
3	0.063	0.090	0.302	0.157	0.098	0.158	0.567	0.305
Big	0.014	0.332	0.053	0.454	0.024	0.602	0.106	0.847

Notes:

For correlation matrix: ** p-value≤0.01, * p-value≤0.05

Table 4.9: Descriptive statistics: 16 portfolios, period 2 (1998-2009)

Explanatory returns

	Sample moments					Correlation matrix			
	Rm	SMB	HML	WML		RM	SMB	HML	WML
Mean	-0.144	-0.250	0.714	0.603	RM	1.000	-0.091	-0.262*	-0.255*
Median	0.052	-0.527	0.880	1.241	SMB	-0.091	1.000	0.085	-0.262*
Maximum	17.924	12.795	7.950	15.058	HML	-0.262*	0.085	1.000	-0.046
Minimum	-20.371	-14.072	-6.663	-25.299	WML	-0.255*	-0.262*	0.046	1.000
t-value	-0.325	-0.807	3.226	1.324					

Dependent returns: portfolios formed on size and book-to-market equity

	Mean				t values			
	Low	2	3	High	Low	2	3	High
Small	-0.399	-0.034	0.086	0.124	-0.585	-0.064	0.176	0.258
2	-0.693	-0.259	0.078	0.197	-1.102	-0.517	0.163	0.371
3	-0.244	-0.026	0.159	0.297	-0.417	-0.055	0.347	0.561
Big	-0.376	0.144	0.627	0.370	-0.799	0.359	1.366	0.722

Dependent returns: portfolios formed on size and momentum

	Mean				t values			
	Loser	2	3	Winner	Loser	2	3	Winner
Small	0.203	0.180	0.334	0.600	0.298	0.369	0.749	1.180
2	-0.389	-0.137	0.161	0.308	-0.270	0.924	0.363	0.597
3	-0.512	-0.166	0.090	0.330	-0.737	-0.324	0.202	0.661
Big	-0.313	-0.387	-0.508	0.219	-0.423	-0.743	-1.126	0.454

Notes:

For correlation matrix: ** p-value$\leq$0.01, * p-value$\leq$0.05

Table 4.10: Differences in means for extreme portfolios: period 1 (1984-1998)

Portfolios formed on size and book-to-market equity

	Difference in means				t values			
	Small	2	3	Big	Small	2	3	Big
High-low	0.534	0.668	0.630	0.700	2.479	3.468	3.924	2.472
	Low	2	3	High	Low	2	3	High
Small-big	0.298	-0.118	0.031	0.132	0.758	-0.312	0.100	0.368

Portfolios formed on size and momentum

	Difference in means				t values			
	Small	2	3	Big	Small	2	3	Big
Winner-Loser	-0.400	-0.077	0.094	0.440	-1.463	-0.300	1.318	1.148
	Loser	2	3	Winner	Loser	2	3	Winner
Small-big	0.659	0.432	0.935	-0.181	2.035	1.231	2.816	-0.479

Table 4.11: Differences in means for extreme portfolios: period 2 (1998-2009)

Portfolios formed on size and book-to-market equity

	Difference in means				t values			
	Small	2	3	Big	Small	2	3	Big
High-low	0.523	0.889	0.541	0.746	2.381	3.670	2.590	2.866
	Low	2	3	High	Low	2	3	High
Small-big	-0.022	-0.178	-0.540	-0.245	0.058	-0.636	-2.218	-0.997

Portfolios formed on size and momentum

	Difference in means				t values			
	Small	2	3	Big	Small	2	3	Big
Winner-Loser	0.387	0.696	0.842	0.531	1.244	2.372	2.690	1.182
	Loser	2	3	Winner	Loser	2	3	Winner
Small-big	0.516	0.567	0.841	0.371	1.531	2.133	3.555	1.161

Table 4.12: Comparison of tests: **a=0**

	Size-value-sorted portfolios		Size-momentum-sorted portfolios	
	Model (2)	Model (3)	Model (2)	Model (3)
Full period	0.279	0.332	9.5e-0.6	6.5e-0.6
Period 1	0.775	0.842	0.016	0.074
Period 2	0.687	0.883	0.038	0.104

Table 4.13: Model (1): 16 portfolios, period 1 and period 2

Model (1): period 1 (1984-1998)

	b_i				$t(b_i)$			
	Low	2	3	High	Low	2	3	High
Small	0.964	0.899	0.865	0.807	9.307	9.695	9.753	9.057
2	0.955	0.967	0.961	0.919	11.259	13.274	10.970	11.382
3	1.043	1.027	0.968	1.025	14.888	15.322	20.422	18.633
Big	1.100	1.054	0.978	0.877	26.620	23.223	31.015	15.348

	a_i				$t(a_i)$			
	Low	2	3	High	Low	2	3	High
Small	-0.000	0.002	0.003	0.005	-0.044	0.423	0.696	1.020
2	-0.004	-0.002	0.001	0.003	-0.886	-0.538	0.250	0.674
3	-0.005	-0.001	0.000	0.001	-2.035	-0.361	0.024	0.262
Big	-0.003	0.003	0.003	0.004	-1.573	1.621	1.900	1.532

Model (1): period 2 (1998-2009)

	b_i				$t(b_i)$			
	Loser	2	3	Winner	Loser	2	3	Winner
Small	0.993	0.789	0.760	0.732	8.391	8.425	8.913	8.353
2	0.953	0.845	0.801	0.863	8.226	9.915	10.175	10.552
3	1.034	0.862	0.824	0.921	11.536	11.678	12.172	11.136
Big	1.045	0.829	0.938	0.882	55.064	17.603	18.105	10.404

	a_i				$t(a_i)$			
	Loser	2	3	Winner	Loser	2	3	Winner
Small	-0.003	0.000	0.002	0.002	-0.553	0.364	3.278	0.426
2	-0.006	-0.002	0.002	0.002	-1.210	-0.479	0.415	0.630
3	-0.001	0.001	0.002	0.004	-0.316	0.214	0.682	0.909
Big	-0.002	0.003	0.008	0.005	-2.289	1.300	3.324	1.354

Linear hypothesis test for **a=0**

	Period 1		Period 2
p-value	0.192	p-value	0.008

Table 4.14: Model (2): 16 portfolios, period 1 and period 2

Model (2): period 1 (1984-1998)

	b_i				$t(b_i)$			
	Low	2	3	High	Low	2	3	High
Small	1.037	0.971	0.948	0.934	29.857	24.300	28.194	24.065
2	0.982	1.041	1.041	1.038	30.237	45.261	55.082	51.168
3	1.061	1.069	1.031	1.108	30.316	28.616	34.448	35.884
Big	1.024	1.043	1.011	0.961	43.108	25.279	39.744	23.558
	s_i				$t(s_i)$			
	Low	2	3	High	Low	2	3	High
Small	1.133	1.062	0.988	0.957	19.537	27.550	15.803	17.505
2	0.981	0.860	0.914	0.870	17.296	28.120	39.486	30.085
3	0.616	0.583	0.572	0.574	12.015	13.348	10.427	11.348
Big	-0.141	-0.233	-0.041	-0.193	-4.173	-3.674	0.861	-3.117
	h_i				$t(h_i)$			
	Low	2	3	High	Low	2	3	High
Small	-0.060	-0.018	0.134	0.540	-0.489	-0.159	1.094	4.060
2	-0.371	0.123	0.151	0.519	-4.279	2.120	3.184	9.059
3	-0.224	0.012	0.206	0.380	-2.289	0.176	2.441	5.513
Big	-0.597	0.051	0.323	0.869	-7.245	0.597	3.619	6.373
	a_i				$t(a_i)$			
	Low	2	3	High	Low	2	3	High
Small	0.001	0.003	0.003	0.002	0.322	1.412	1.310	1.459
2	-0.001	-0.002	0.001	0.000	-0.461	-1.414	0.510	0.378
3	-0.004	-0.001	-0.001	-0.001	-1.844	-0.410	-0.441	-0.815
Big	0.000	0.002	0.001	-0.001	0.001	1.565	0.479	-0.661

Model (2): period 2 (1998-2009)

	b_i				$t(b_i)$			
	Low	2	3	High	Low	2	3	High
Small	1.094	0.907	0.868	0.870	16.531	19.180	14.598	18.610
2	1.014	0.932	0.913	1.044	19.226	22.592	24.556	50.399
3	1.078	0.956	0.943	1.081	23.529	17.630	19.735	22.464
Big	1.015	0.865	1.001	1.012	68.083	18.255	21.354	15.747
	s_i				$t(s_i)$			
	Low	2	3	High	Low	2	3	High
Small	1.467	1.095	0.954	0.931	17.046	17.414	14.638	15.882
2	1.372	0.977	0.870	0.918	20.907	20.353	24.800	28.810
3	0.998	0.663	0.565	0.677	15.120	12.286	11.637	14.873
Big	-0.126	-0.116	0.062	0.097	-4.199	-3.176	0.924	1.131
	h_i				$t(h_i)$			
	Low	2	3	High	Low	2	3	High
Small	0.107	0.404	0.391	0.633	0.582	3.069	2.464	5.692
2	-0.153	0.218	0.465	0.965	-1.083	2.350	4.128	10.867
3	-0.118	0.420	0.658	0.918	-0.998	3.496	5.619	7.363
Big	-0.170	0.328	0.449	0.950	-3.216	3.013	4.694	5.188
	a_i				$t(a_i)$			
	Low	2	3	High	Low	2	3	High
Small	0.001	0.001	0.002	0.000	0.140	0.225	0.572	0.131
2	-0.001	-0.000	0.001	-0.001	-0.266	-0.185	0.390	-0.818
3	0.002	-0.000	-0.000	-0.000	0.864	-0.097	-0.195	-0.169
Big	-0.001	0.000	0.005	-0.001	-1.292	0.018	1.945	-0.330

Linear hypothesis test for **a=0**

p-value	0.775	p-value	0.687

Table 4.15: Model (3): 16 portfolios, period 1 and period 2

Model (3): period 1 (1984-1998)

	b_i				$t(b_i)$			
	Loser	2	3	Winner	Loser	2	3	Winner
Small	0.995	1.003	1.006	0.944	26.948	33.552	25.650	20.959
2	1.015	1.040	1.052	0.967	37.249	50.240	38.607	33.531
3	1.064	1.059	1.060	1.024	28.401	28.552	24.254	24.325
Big	0.977	0.991	1.004	1.047	26.736	18.004	21.999	30.067
	s_i				$t(s_i)$			
	Loser	2	3	Winner	Loser	2	3	Winner
Small	1.027	1.068	1.040	1.006	20.015	21.644	18.689	17.741
2	0.805	0.831	0.815	0.878	27.613	23.611	39.486	20.705
3	0.530	0.557	0.540	0.540	10.595	12.200	10.153	9.767
Big	-0.071	-0.184	-0.119	-0.143	-1.247	-2.232	-2.034	-3.117
	h_i				$t(h_i)$			
	Loser	2	3	Winner	Loser	2	3	Winner
Small	0.157	0.127	0.165	0.021	1.658	1.417	1.619	0.183
2	0.206	0.149	0.166	0.093	3.484	2.008	2.638	1.181
3	0.211	0.227	0.177	0.009	1.720	3.270	2.127	0.096
Big	0.069	0.119	0.398	0.126	0.438	0.647	3.578	1.309
	m_i				$t(m_i)$			
	Loser	2	3	Winner	Loser	2	3	Winner
Small	-0.454	-0.147	0.169	0.328	-8.953	-3.334	3.111	4.635
2	-0.528	-0.212	0.053	0.323	-16.023	-6.085	1.137	6.628
3	-0.676	-0.320	0.010	0.327	-12.859	-7.637	0.157	5.021
Big	-0.845	-0.475	-0.089	0.450	-9.289	-5.089	-1.259	6.146
	a_i				$t(a_i)$			
	Loser	2	3	Winner	Loser	2	3	Winner
Small	0.004	0.006	0.007	0.001	2.920	2.864	3.391	0.374
2	-0.001	0.001	0.001	-0.002	-0.419	0.983	0.812	-0.800
3	-0.002	-0.002	0.000	-0.001	-0.795	-1.234	0.024	-0.395
Big	-0.002	0.001	-0.004	0.001	-0.587	0.238	-1.787	0.426

Model (3): period 2 (1998-2009)

	b_i				$t(b_i)$			
	Loser	2	3	Winner	Loser	2	3	Winner
Small	0.945	0.771	0.795	0.887	16.495	13.209	9.046	9.105
2	1.002	0.863	0.846	0.970	20.015	14.509	11.615	10.648
3	1.066	0.891	0.876	1.037	18.057	14.115	10.748	11.062
Big	1.056	0.914	0.887	1.063	13.059	11.926	13.150	24.275
	s_i				$t(s_i)$			
	Loser	2	3	Winner	Loser	2	3	Winner
Small	1.223	0.946	0.889	1.037	13.459	15.978	9.530	7.246
2	1.031	0.794	0.795	1.012	14.618	14.046	11.100	9.970
3	0.772	0.593	0.580	0.774	11.835	8.175	8.087	8.902
Big	-0.015	-0.015	0.019	0.088	-0.127	-0.185	0.260	1.236
	h_i				$t(h_i)$			
	Loser	2	3	Winner	Loser	2	3	Winner
Small	0.285	0.431	0.438	0.319	2.108	4.759	3.607	1.961
2	0.252	0.514	0.483	0.359	2.067	5.090	3.882	2.000
3	0.334	0.554	0.514	0.346	2.632	4.680	3.629	2.300
Big	0.017	0.051	0.194	-0.044	0.137	0.352	1.288	-0.480
	m_i				$t(m_i)$			
	Loser	2	3	Winner	Loser	2	3	Winner
Small	-0.504	-0.197	0.045	0.271	-7.439	-3.218	0.494	2.136
2	-0.557	-0.267	0.016	0.245	-9.250	-4.766	0.205	2.065
3	-0.648	-0.313	-0.062	0.277	-9.409	-4.368	-0.753	2.641
Big	-0.949	-0.422	-0.194	0.425	-9.968	-7.077	-2.153	5.970
	a_i				$t(a_i)$			
	Loser	2	3	Winner	Loser	2	3	Winner
Small	0.007	0.003	0.003	0.005	2.478	1.326	1.004	1.237
2	0.001	0.000	0.001	0.003	0.659	-0.099	0.419	0.779
3	-0.000	-0.001	0.000	0.003	-0.057	-0.390	0.066	0.794
Big	0.004	-0.000	-0.004	0.002	1.356	-0.117	-1.404	0.755

Linear hypothesis test for **a=0**

	Period 1		Period 2
p-value	0.074	p-value	0.104

Chapter 5

Addressing Biases in Dynamic Linear Panel Models: An Implementation in R with Monte Carlo Evidence for Growth Regressions

As it has been shown by Nickell (1981) the Least square dummy variable estimator (LSDV) is not consistent for large N and finite T in dynamic linear panel data models. The growth regression frequently used in macroeconomic research provides a prominent application in this regard. Based on a statistical simulation analysis we evaluate the performance of the bias correction procedure originally proposed by Kiviet (1995) tackling the consistency problem in dynamic linear panel models in the context of growth empirics. Additionally, we provide a baseline framework for a generally accessible function which computes this bias-corrected LSDV estimator in R. Simulation results show that bias performance of the convergence parameter is improved considerably by the bias correction procedure. However, is does not reduce the average bias over all parameters.

5.1 Introduction

Understanding the process of economic growth is a crucial subject in macroeconomics. For the empirical evaluation of observed cross-country income differences dynamic models for panel data have gained increasing interest among economic researchers. The econometric foundation has been provided by Islam (1995), which reformulates the growth regression for a dynamic panel data setting with fixed effects. Whereas different statistical methodologies might be applied for parameter identification, the Least square dummy variable estimator (LSDV)[1] is a preferred estimator in the empirical growth literature as it controls for unobserved country-specific heterogeneity.

A well-known problem in the context of dynamic panel data models is that the LSDV estimator is not consistent for large N and finite T (Nickell, 1981). Two main classes of solutions have been proposed to overcome this inconsistency problem. The first class consists of IV and GMM estimation techniques applying instruments to consistently identify the model parameters (see e.g. Arellano and Bond (1991)). The second class consists of bias-corrected LSDV estimators, where the small sample bias of the LSDV estimator is numerically approximated (see e.g. Kiviet, 1995). Looking at the literature it is to note that the first type of estimation procedures is commonly used in practice, whereas the latter does not often appear in applied economic research. Furthermore, compared to the usual panel data estimation methodologies, bias-corrected LSDV estimators are less frequently implemented for software commonly used by economic researchers and practitioners.

As a special case of a dynamic linear panel model, the growth regression is subject to this consistency problem referred to as Nickell-Bias. Different studies examine statistical properties of commonly used estimators including the IV and GMM approaches.[2] To our knowledge, no study has compara-

[1] As already noted in chapter 1, the LSDV estimator is generally considered one version of the fixed effects approach. In the remainder of this chapter I will use the LSDV notation.

[2] See Hauk and Wacziarg (2009) or Bond et al. (2001).

tively analysed the class of bias-correction estimators within the context of economic growth. Given the increasing interest in understanding observed aggregate income differences and the according concern of unjustified claims of causality in parameter identification, it appears to be rewarding to evaluate estimation procedures in growth empirics in order to guide researchers towards less biased estimations. Using simulation methodology, we therefore evaluate the bias properties of the bias corrected LSDV estimator relative to most commonly used panel data estimators for growth regressions. Hence, the main contribution of this paper is the inclusion and evaluation of the described bias-correction procedure in the context of growth empirics. We additionally provide a baseline framework for a generally accessible function which computes the bias-corrected LSDV estimator in R.

Different studies have applied simulation methodology in order to evaluate the bias-corrected LSDV estimator. However, these analysis have been conducted within a general setting. Kiviet (1995) shows that by using a general AR(1) process with one exogenous variable the bias-corrected LSDV estimator outperforms the usual IV and GMM estimators for varying sizes of T. By using similar simulation techniques, Judson and Owen (1999), Bun and Carree (2005) and Bruno (2005) all provide evidence for the superiority of the bias-corrected LSDV estimator for dynamic panel-data models with finite N. Our simulation analyses show that when we apply the bias-corrected estimators to growth regressions, superiority cannot be maintained. Although the bias correction provides an improvement on the identification of the lagged dependent variable (convergence parameter), it is outperformed by other estimators in terms of overall bias.

The remainder of the paper is organized as follows. In a first part (section 5.2 and 5.3) we give an outline of the basic statistical problem related to the general first order linear dynamic panel data model. Accordingly, we sketch the bias-approximations which will be used for computational implementation. In a second part (section 5.4) we provide a simulation analysis by applying the bias correction procedure to a data-generating process representing economic growth and evaluating its bias properties relative to other commonly used panel estimators. Section 5.5 summarizes and concludes.

5.2 The Baseline Problem

We consider the standard first order linear dynamic panel data model with K exogenous explanatory variables x_{it}:

$$y_{it} = \gamma y_{i,t-1} + \beta' x_{it} + \eta_i + \epsilon_{it}, \quad i = 1, 2, ..., N; \; t = 1, 2, ..., T. \tag{5.1}$$

The dependent variable y_{it} is regressed on the vector x_{it} and on its one period lagged value $y_{i,t-1}$. $\eta_i + \epsilon_{it}$ is the composite disturbance with η_i being the unobserved individual specific effect and ϵ_{it} being an unobserved white noise error term with constant variance σ^2_ϵ.
Stacking the observations over time and across individuals we can rewrite the model

$$y = W\delta + (I_N \otimes \iota_T)\eta + \epsilon, \tag{5.2}$$

where $\delta = (\gamma, \beta')'$, y and $W = (y_{-1}, X)$ are $NT \times 1$ and $NT \times (K+1)$ matrices of stacked observations, ϵ is the $NT \times 1$ vector of disturbances and $\iota_T = (1, ..., 1)'$ a $T \times 1$ vector of ones. The LSDV estimator is given by

$$\hat{\delta}_{LSDV} = (W'AW)^{-1}W'Ay. \tag{5.3}$$

$A = I_N \otimes (I_T - (1/T)\iota_T\iota_T')$ is a $NT \times NT$ transformation matrix which eliminates the individual specific effects.
The consistent estimation of $\hat{\delta}_{LSDV}$ requires the strict exogeneity assumption to hold, $E(\epsilon_{it}|x_i, y_{i,t-1}, \eta_i) = 0$ for $t = 1, 2, ...T; i = 1, 2, ..., N$. It is well established in the panel data literature that in the usual fixed effects setting with $N \to \infty$ and finite T the inclusion of a lagged dependent variable y_{-1} in W violates this assumption. More specifically, Nickell (1981) examines the bias of γ when there are no exogenous regressors and shows that it approaches

zero as $T \to \infty$.

Accordingly, several estimators have been proposed for applications with small T. Standard approaches are IV and GMM estimation techniques exploiting the panel structure of the data to construct valid instruments (see Anderson and Hsiao, 1981; Arellano and Bond, 1991; or Blundell and Bond, 1998). An alternative but less common strategy in practice is the additive correction of the inconsistent estimator $\hat{\delta}_{LSDV}$ by a numerical approximation of the bias (see Kiviet, 1995; Hansen, 2001; Hahn and Kuersteiner, 2002; Bruno, 2005; or Bun and Carree, 2005). The following three bias-approximation terms - $B1$, $B2$ and $B3$ - derived by Bun and Kiviet (2003) and based on Kiviet (1995) will be used for the computational implementations and simulation analysis in the next section (with an increasing level of accuracy):

$$B1 = c1; \quad B2 = B1 + c2; \quad B3 = B2 + c3, \tag{5.4}$$

with

$$c1 = \sigma_\epsilon^2 tr(\Pi) q_1, \tag{5.5}$$

$$c2 = -\sigma_\epsilon^2 [Q\bar{W}\Pi A\bar{W} + tr(Q\bar{W}'\Pi A\bar{W}) I_{K+1} + 2\sigma_\epsilon^2 q_{11} tr(\Pi'\Pi\Pi) I_{K+1}] q_1 \tag{5.6}$$

$$c3 = \sigma_\epsilon^4 tr(\Pi)\{2q_{11}Q\bar{W}'\Pi\Pi'\bar{W}q_1 + [(q_1'\bar{W}'\Pi\Pi'\bar{W}q_1) + q_{11}tr(Q\bar{W}'\Pi\Pi'\bar{W}) + 2tr(\Pi'\Pi\Pi'\Pi)q_{11}^2]q_1\}, \tag{5.7}$$

and

$Q = [E(W'AW]^{-1} = [\bar{W}'A\bar{W} + -\sigma_\epsilon^2 tr(\Pi'\Pi)e_1 e_1']^{-1}$,

$\bar{W} = E(W)$,

$e_1 = (1, 0, ..., 0)'$,

$\Pi = AL\Gamma, \Gamma = I_N \otimes \Gamma_T, \Gamma_T = (I_T - \gamma L_T)^{-1}, L = I_N \otimes L_T$,

where e_1 has $K+1$ elements. The $T \times T$ matrix L_T has ones on its first lower subdiagonal and all other elements equal zero. Furthermore, $q_1 = Qe_1$ and $q_{11} = e_1' q_1$.
The approximation terms will be used for computational implementation described in the next section.

5.3 Implementing the Bias-Corrected LSDV Estimator in R

Bruno (2005) provides a computational implementation of the bias-corrected LSDV estimator which is available as a Stata routine *xtlsdvc*. The same estimator does not exist as a software application in R, a free environment for statistical computing. As R is increasingly used by economic researchers and practitioners it appears to be rewarding to provide statistical procedures which are frequently used in the field (see Kleiber and Zeileis, 2008). Accordingly, we provide a function which computes the bias-corrected LSDV within the R environment using the derivations described in the last section.[3]
The bias-corrected LSDV estimator is obtained by subtracting the approximation terms presented in the last section from the LSDV estimator $\hat{\delta}_{LSDV}$. Since the bias approximation is evaluated at the unobserved true parameters, we need corresponding values for γ and σ_ϵ in a first step in order to calculate $\hat{B}_i$. Consistent estimators can be obtained by using the GMM procedures described above. In the following we will use the GMM estimator for the first-differenced model described by Arellano and Bond (1991) (AB) to calculate $\hat{\gamma}$ and $\hat{\sigma_\epsilon}$. The bias-corrected LSDV estimators are

$$\hat{\delta}_{LSDVC_i} = \hat{\delta}_{LSDV} - \hat{B}_i, \qquad i = 1, 2, 3. \tag{5.8}$$

[3]The corresponding file containing the R-procedure (LsdvcR.txt) can be obtained by the author upon request. The implementation of a publicly accessible R software package is work in progress.

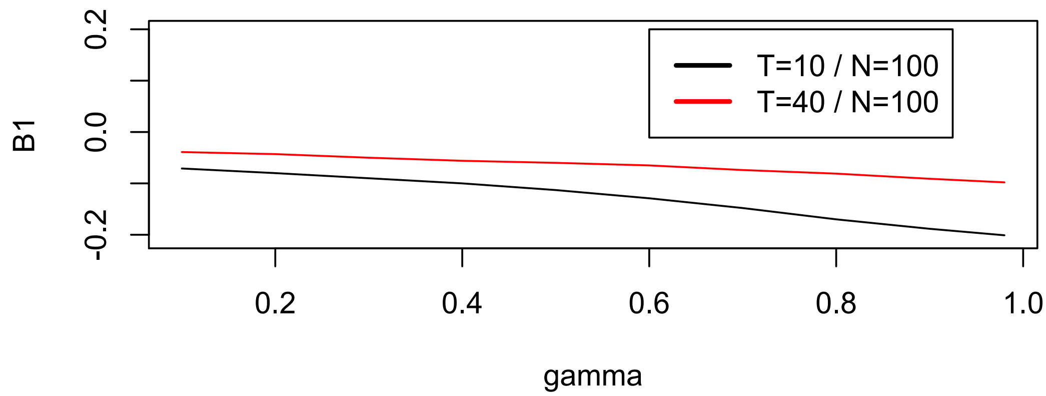
We generate one data sample using the process described by model (5.1) with one exogenous regressor x_{it} in order to explore the basic properties of the estimator computed in R.[4] The parameter values for data generation are $\gamma = 0.8$ and $\beta = 0.8$. Table 5.1 gives the estimators $\hat{\gamma}$ and $\hat{\beta}$ computed by the implemented R-procedure for the LSDV estimator $\hat{\delta}_{LSDV}$, the bias-corrected estimators $\hat{\delta}_{LSDVC_i}$ as well as the first step AB estimator $\hat{\delta}_{AB}$. The identical estimated values can be obtained by using the Stata routine *xtlsdvc*. As already shown by other studies (see e.g. Bun and Kiviet, 2002), compared to the true parameters, the LSDV estimator bias the parameter downwards with $\hat{\gamma} = 0.620$ and $\hat{\beta} = 0.722$. By applying the bias correction procedure the estimated parameter values increase towards the true coefficients decreasing the extent of the bias. Varying parameter values in the data-generating process we can reproduce some well-documented properties about this estimator (see e.g. Judson and Owen, 1999): Figure 5.1 shows that the bias $B1$ is always smaller (in absolute terms) with a higher T and that the bias increases (in absolute terms) with a higher γ.

Table 5.1: Computing LSDVC estimates in R

	δ (True coefficients)	$\hat{\delta}_{LSDV}$	$\hat{\delta}_{AB}$ (First stage)	$\hat{\delta}_{LSDVC_1}$	$\hat{\delta}_{LSDVC_2}$	$\hat{\delta}_{LSDVC_3}$
$y_{i,t-1}$	0.800	0.620	0.686	0.736	0.736	0.736
x_{it}	0.800	0.722	0.727	0.749	0.748	0.748

[4]Data are generated using the Stata routine *xtarsim*. Detailed description is given by Bruno (2005).

Figure 5.1: LSDV bias estimates



5.4 An Application to Growth Empirics

In this section we evaluate the bias properties of the corrected LSDV estimator for growth regressions using Monte Carlo simulations. Following the seminal works by Baumol (1986) and Barro (1991) the growth regression has become the standard model for empirical evaluations of cross country aggregate income differences. Based on the theoretical work by Solow (1956) and the according statistical specification derived by Mankiw et al. (1992), Islam (1995) reformulates the growth regression for a panel data setting. The functional form of this model can be represented as

$$\log y_{it} = \theta' x_{it} + \eta_i + \mu_t + \upsilon_{it}, \tag{5.9}$$

where t denotes the end of a time period of duration τ and $t-\tau$ is the beginning of this period. We define $x'_{it} = [1, \log s_{k,it-\tau}, \log(n+g+\rho), \log y_{it-\tau}]$ and $\theta' = [\theta_0, \theta_1, \theta_2, \theta_3]$. s_k is the country's saving rate in physical capital, n and g are the growth rates of the country's population and technology level, and ρ is the rate at which these variables depreciate. θ' is the vector including the reduced-form parameters.[5] The regression equation (5.9) represents a special case of a dynamic linear panel model with a moderate T. Typically, for growth regression the number of observations over time is between 5 and 7. Focusing on the long-run growth effects, data are usually averaged over 5 years lag ($\tau = 5$) in order to exclude short-run business cycle effects.

In addition to the bias caused by the endogeneity of the lagged dependent variable, the typical sources of bias in the context of growth empirics are omitted country specific effect, reverse causality in the regressors, and the regressors measurement errors. By performing Monte Carlo simulations Hauk and Wacziarg (2009) consider the commonly used estimators for growth regressions and evaluate the according bias properties. They find that depending on the specific source of bias, different classes of estimators perform best. Generally, throughout different simulation designs, the between estimator (OLS applied to a single cross-section averaged over time) reveals a relatively stable estimation performance especially with regard to income convergence. We extend the simulation analysis performed by Hauk and Wacziarg (2009) by allowing for a bias correction in the LSDV estimator.

5.4.1 Simulation Methodology

We apply the simulation analysis suggested by Hauk and Wacziarg (2009) in order to compare the performance of the bias corrected LSDV estimator (LSDVC) with the bias properties of the most commonly used econometric methods for growth regressions. Apart from the uncorrected LSDV estimator and two GMM procedures (Arellano-Bond GMM (AB) and Blundell-Bond GMM (BB))[6] we additionally evaluate bias properties of the OLS, the Random effects (RE) and the Between estimator (BE). The latter is specifically

[5]A detailed derivation of the growth regression as well as a representation of the implied structural parameters can be found in Mankiw et al. (1992) or in Islam (1995).

[6]The BB estimator is also referred to as SYS-GMM estimator. I will use the former term in this chapter.

interesting for exploring estimation performance when focusing on the among-group component of total variation.

We obtain the underlying data from the Penn World Tables version 7.1 (PWT 7.1): $\log s_k$ is captured by using the log of investment rates as share of GDP, n is the rate of population growth, and $\log y_{it}$ is the log of per capita income in purchasing power parity. Following the usual convention we assumed that $g + \rho = 0.07$. We define the interval τ to be 4 years and underlying data spans from 1973 to 2007 giving $T = 7$ observations over time. In order to have a balanced panel the number of countries is reduced to $N = 80$.

Model (5.9) represents our data generating process. The variables in $x'_{it} = [1, \log s_{k,it-\tau}, \log(n + g + \rho), \log y_{it-\tau}]$ are simulated by using moments of the corresponding observed variables obtained from PWT 7.1. We simulate the country-fixed effect η_i by running a fixed-effects regression on the observed data. We array all observations in a $N \times (T(K-1)+2)$ matrix. From these N observations we compute the $(T(K-1)+2) \times 1$ vector of means for these variables, $\hat{m}_{\theta,\eta}$, as well as their variance covariance matrix $\hat{\Omega}_{\theta,\eta}$. Based on these estimates we define a mulitvariate normal distribution process which is used to draw N observations for each run of the simulation.

We consider the possible endogeneity of the regressors $\log s_{k,it-\tau}$ and $\log(n + g + \rho)$ by simulating the residuals υ_{it}. We allow the residual term to be correlated with the corresponding variables. Following economic consideration we expect a higher income to encourage investment and decrease reproduction activities. Correspondingly, $\log s_k$ is positively and $\log(n + g + \rho)$ negatively correlated with υ_{it}. The residuals are additionally allowed to be correlated over time. We obtain the underlying values for υ_{it} by running a fixed-effects regression using the observed values on model (9). We then array them in a $N \times T$ matrix and compute $T \times T$ variance covariance matrix $\hat{\Omega}_{\upsilon}$. The simulated variables are computed by drawing N values from a corresponding multivariate normal distribution with mean 0. The parameter values combined in the vector θ are chosen conventionally with $\theta_1 = 0.3$, $\theta_2 = -0.197$,

and $\theta_3 = 0.832$.[7]
The simulated dependent variable for period 2, y_{i2}, is computed by applying the data-generating process (5.9) using all parameters and simulated variables described above. The computed value for y_{i2} is then used to analogously generate y_{i3}. This process is repeated iteratively up to y_{iT}.

5.4.2 Simulation Results

Table 5.2 presents the computed estimates based on averages calculated over 1000 runs. We vary the extent of the endogeneity bias by varying the residual correlation with the regressors between 0% and 50%. The difference between the averages of the estimates and the corresponding true parameters gives a measure of the according biases.[8]
Considering the average absolute bias over all coefficients our results show that the RE and BB estimators perform best with the corresponding measure lying between 15% and 22 %. Also, in terms of average absolute bias, the BE estimator performs better then the LSDV estimator suggesting that focussing only on the within-variation increases the bias. The relatively high average absolute bias of the AB estimator (around 80%) indicates the well-known problem of weak instruments arising in small samples (small N) (see Stock et al., 2002). Accordingly, adding more instruments for the BB system GMM procedure improves estimation properties considerably resulting in a relatively moderate bias ranging around 20 %.
Looking at the average absolute bias, the bias-corrected LSDV estimator does not outperform its LSDV counterpart. The corresponding measure is about 20% higher then without the correction procedure. Considering the individ-

[7] For a detailed discussion on the economic implications with regard to the underlying sturctural parameters in the context of the Solow model see for example Barro and Sala-i-Martin (1995).

[8] We only focus on the biases and do not evaluate the estimator's standard errors and efficiency properties.

ual coefficients we see that it is especially the parameter for $\log(n+g+\rho)_{it-\tau}$, θ_2, which exhibits a relatively high bias (around 135 %). This problem may arise from plugging in the coefficients estimated by the GMM method in the first stage of the bias-correction procedure.[9] As we see from table 5.2 the AB estimations of θ_2 exhibits a relatively high bias which is due to the weak instrument problem. However, comparing the LSDV and the LSDVC estimators, we find that the convergence parameter[10] θ_3 is considerably smaller after the bias correction with relative biases ranging around 0.9%. In fact, with regard to the convergence parameter, the LSDVC estimator reveals the best performance throughout all estimation methods.

The signs of the biases correspond mostly to previous findings (see Hauk and Wacziarg (2009)) : the BE estimator tends to bias the parameter θ_3 upwards, whereas the LSDV, LSDVC and AB estimators reveal a negative bias implying a higher speed of convergence.

The bias properties of θ_1, the slope parameter for $\log s_{k,it-\tau}$, also varies across estimators in sign and magnitude. Due to instrumentation, the bias is small for the BB estimators (around 6%) whereas it appears to be relatively high for the LSDV, LSDVC and AB estimator. OLS, BE and RE exhibit a moderate bias for θ_1 ranging form 15% to 35%. The estimators are biased downward throughout. θ_2, the slope parameter for $\log(n+g+\rho)_{it-\tau}$, is moderate for the RE estimator (around 10%) but exhibits relatively high biases for all other methods with all relative measures being higher then 60%. As mentioned above, the difference between the true and the estimated coefficients is especially high for the AB estimator. The sign is positive for BE and RE, exploiting between variation, whereas the LSDV, LSDVC, and AB tend to underestimate the parameters.

Varying the extent of endogeneity between 0% and 50% in columns (1) to (4)

[9] The AB as well as the BB estimator used for the first step estimation lead to a relative high average bias for the LSDVC estimator.

[10] Following the usual notation in the empirical growth literature we refer to the parameter of the lagged dependent variable as convergence parameter.

reveals that, in terms of average absolute biases, results remain mostly unaffected. Expectedly, the GMM estimators using instruments to overcome the endogeneity problem slightly improve their relative performance. However, the effect is not substantial. According to Hauk and Wacziarg (2009), this effect may arise from the relatvely low variance of v_{it} compared to the regressors, translating into a small covariance and therefore a small endogeneity bias.

The superior performance of the RE estimator is rather surprising. Especially in the field of applied Macroeconomics, the LSDV estimator is usually preferred over the RE model since it controls for country specific heterogeneity. It is expected that these country specific characteristics contained in the individual specific effect are correlated with the other regressors violating the exogeneity assumption. Our findings show that in finite samples, for data representing the growth process, the LSDV estimator which focuses on the within-variation only reveals generally higher biases. The bias correction procedure improves bias properties for the convergence parameter, however, it does not improve the overall bias.

5.5 Conclusions

Using simulation methodology we analyse bias properties of the bias corrected LSDV estimator proposed by Kiviet (1995) for growth regressions and compare it to commonly used panel data estimation methodologies. We specifically focus on the finite sample bias caused by the dynamic specification of the model. We contribute to the existing literature by including the bias-correction procedure into the field of empirical growth analysis and by providing a baseline implementation of a function which computes this bias-corrected LSDV estimator in R.

The bias corrected LSDV estimator shows superior bias properties for the estimated slope parameter for the lagged dependent variable. Thus, in the context of growth regressions, the correction procedure helps to identify the convergence parameter more precisely.

Simulation results further suggest that retaining between variation for estimation reduces the overall bias. Accordingly, compared to LSDV, the averaged biases are smaller for the BE and RE. Also, comparing the AB and its extended version the BB estimator, we see that the latter performs considerably better indicating to the weak instrument problem in small samples for the AB estimator.

Varying the extent of the endogeneity bias does not considerably changes the overall performance of the single estimators. All in all, the RE estimator performs best in terms of overall bias in finite sample. This result may be driven by the fact, that the estimation procedure exploits a combination of within- and between- country variation. Also, the performance of the BB estimator is relatively good, which is due to the application of valid instrument.

Table 5.2: Average estimated coefficients and biases (1000 runs)

	True coefficient	(1)		(2)		(3)		(4)	
Error corr.		0%		25%		40%		50%	
		Avg. coef.	Bias (%)	Avg. coef.	Bias (%)	Avg. coef.	Bias (%)	Avg. coef.	Bias (%)
OLS									
$\log s_{k,it-\tau}$	0.3	0.254	-15.402	0.253	-15.593	0.254	-15.446	0.255	-15.125
$\log(n+g+\rho)_{it-\tau}$	-0.197	-0.329	67.02	-0.332	68.509	-0.332	68.739	-0.324	64.409
$\log y_{it-\tau}$	0.832	0.878	5.508	0.878	5.565	0.878	5.540	0.878	5.547
Avg. Bias			29.310		29.889		29.908		28.360
BE									
$\log s_{k,it-\tau}$	0.3	0.217	-27.728	0.217	-27.744	0.217	-27.540	0.217	-27.520
$\log(n+g+\rho)_{it-\tau}$	-0.197	-0.321	62.712	-0.325	64.989	-0.323	64.060	-0.315	60.028
$\log y_{it-\tau}$	0.832	0.91	9.335	0.91	9.387	0.91	9.365	0.91	9.416
Avg. Bias			33.265		34.040		33.655		32.321
LSDV									
$\log s_{k,it-\tau}$	0.3	0.106	-64.527	0.108	-64.049	0.107	-64.224	0.109	-63.668
$\log(n+g+\rho)_{it-\tau}$	-0.197	-0.042	-78.592	-0.046	-76.581	-0.052	-73.369	-0.054	-72.790
$\log y_{it-\tau}$	0.832	0.79	-4.994	0.791	-4.900	0.791	-4.889	0.791	-4.914
Avg. Bias			49.371		48.510		47.494		47.124
RE									
$\log s_{k,it-\tau}$	0.3	0.195	-35.146	0.195	-35.017	0.195	-34.960	0.197	-34.378
$\log(n+g+\rho)_{it-\tau}$	-0.197	-0.214	8.749	-0.218	10.468	-0.223	12.944	-0.220	11.878
$\log y_{it-\tau}$	0.832	0.821	-1.338	0.822	-1.241	0.822	-1.240	0.822	-1.238
Avg. Bias			15.078		15.575		16.381		15.831
AB									
$\log s_{k,it-\tau}$	0.3	-0.006	-101.988	-0.003	-101.013	-0.005	-101.537	-0.004	-101.384
$\log(n+g+\rho)_{it-\tau}$	-0.197	0.065	-132.878	0.063	-132.231	0.056	-128.600	0.051	-126.001
$\log y_{it-\tau}$	0.832	0.804	-3.419	0.805	-3.288	0.804	-3.320	0.804	-3.240
Avg. Bias			79.428		78.844		77.819		76.875
BB									
$\log s_{k,it-\tau}$	0.3	0.281	-6.418	0.28	-6.646	0.281	-6.485	0.281	-6.298
$\log(n+g+\rho)_{it-\tau}$	-0.197	-0.307	55.911	-0.306	55.294	-0.306	55.583	-0.306	55.229
$\log y_{it-\tau}$	0.832	0.857	2.946	0.857	3.005	0.857	2.975	0.857	2.980
Avg. Bias			21.758		21.648		21.681		21.502
LSDVC									
$\log s_{k,it-\tau}$	0.3	0.117	-61.155	0.117	-60.948	0.115	-61.803	0.117	-60.994
$\log(n+g+\rho)_{it-\tau}$	-0.197	0.094	-147.965	0.089	-144.994	0.094	-147.676	0.095	-148.185
$\log y_{it-\tau}$	0.832	0.824	-0.952	0.823	-1.024	0.832	-0.909	0.824	-0.91
Avg. Bias			70.024		67.275		68.989		70.021

References

Acemoglu, D., Johnson, S., Robinson, J.A. (2001). The Colonial Origins of Comparative Development: An Empiricial Investigation. *American Economic Review*, 91(5), 1369-1401.

Akarca, A.T., Long, T.V. (1979). Energy and Employment: a Time Series Analysis of the Causal Relationship. *Resources and Energy*, 2, 151-162.

Akarca, A.T., Long, T.V. (1980). On the Relationship Between Energy and GDP: a Re-Examination. *Journal of Energy and Development*, 5, 326-331.

Alquist, R., Kilian, L. (2007). What Do We learn From the Price of Crude Oil Futures? *Journal of Applied Econometrics* 25, 539-573.

Anderson, T.W., Hsiao, C. (1981). Formulation and Estimation of Dynamic Models Using Panel Data. *Journal of Econometrics*, 18, 47-82.

Andrews, D.W.K., Ploberger, W. (1994). Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative. *Econometrica*, 62, 1383-1414.

Angrist, J.D., Pischke, J.S. (2010). The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Ecnonmetrics. *Journal of Economic Perspectives*, 24 (2), 3-30.

Arellano, M., Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58, 127-134.

Aretz, K., Bartram, S.M., Pope, P.F. (2010). Macroeconomic Risks and Characteristic-Based Factor Models. *Journal of Banking and Finance*, 34/6, 1383-1399.

Bai, J., Perron, P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica*, 66, 47-68.

Bai, J., Perron, P. (2003). Computation and Analysis of Multiple Structural Change Models. *Journal of Applied Econometrics*, 18, 1-22.

Barnes, D., Floor, W.M. (1996). Rural Energy in Developing Countries: A Challenge for Economic Development. *Annual Review of Energy and Environment*, 21, 497-530.

Barro, R. (1991). Economic Growth in a Cross-Section of Countries. *Quarterly Journal of Economics*, 106 (2), 407-443.

Barro, R., Sala-i-Martin, X. (1995). *Economic Growth*. New York: McGraw Hill.

Barsky, R.B., Kilian, L.(2002). Do We Really Know that Oil Caused the Great Stagflation? A Monetary Alternative. In: *NBER, Macroeconomic Annuals 2001* 16, ed. Ben S. Bernanke and Kenneth Rogoff, 27-83, Cambridge: MIT Press.

Baumol, W. (1986). Productivity Growth, Convergence and Welfare: What the Long Run Data Show. *American Economic Review*, 76, 1072-1085.

Bernanke, B. (1986). Alternative Explanations of the Money-Income Correlation. Carnegie-Rochester Conference Series on Public Policy, Amsterdam: North-Holland.

Bernanke, B., Blinder, A. (1992). The Federal Funds Rate and the Channels of Monetary Transmission. *American Economic Review*, 82, 901-921.

Blanchard, O.J., Quah, D. (1989). The Dynamic Effects of Aggregate Demand and Supply Disturbances. *American Economic Review*, 79, 655-673.

Blundell, R, Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Model. *Journal of Econometrics*, 87, 115-143.

Bond, S., Hoeffler, A., Temple, J. (2001). GMM Estimation of Empirical Growth Models. Economic Papers 2001-W21, University of Oxford.

Bouakez, H., Essid, B., Normandin, M., (2013). Stock Returns and Monetary Policy: Are there any Ties? *Journal of Macroeconomics*, 36, 33-50.

Bretschger, L. (2010). Energy Prices, Growth, and the Channels in Between: Theory and Evidence. Economics Working Paper Series 06/47, ETH Zurich.

Büyüksahin, B., Robe, M.A. (2011). Does Oil Matter? Energy Market Financialization and Equity Energy Linkages. Working Paper.

Büyüksahin, B., Haigh, M.S., Harris, J.H., Overdahl, J.A., Robe, M.A. (2008). Fundamentals, trader activity and derivative pricing. *EFA Bergen Meetings Paper. Available at SSRN: ohttp://ssrn.com/abstract=9666924.*

Breitung, J., Brüggemann, R., Lütkepohl, H. (2004). Structural Vector Autoregressive Modeling and Impulse Responses. In: H. Lütkepohl, M. Krätzig (eds.), Applied Time Series Econometrics. Chapter 4, pp. 159-196. Cambridge University Press. Cambridge.

Bruno, G.S.F. (2005). Approximating the Bias of the LSDV Estimator for Dynamic Unbalanced Panel Data Models. *Economics Letters*, 87, 361-365.

Bruno, G.S.F. (2005). Estimation and Inference in Dynamic Unbalanced Panel-Data Models with a Small Number of Individuals. *The Stata Journal*, 5, 473-500.

Bun, M.J.G., Carree, M.A. (2006). Bias-Corrected Estimation in Dynamic Panel Data Models. *Economics Letters*, 92, 220-227.

Bun, M.J.G., Kiviet, J.F. (2003). On the Diminishing Returns of Higher Order Terms in Asymptotic Expansions of Bias. *Economics Letters*, 79, 145-152.

Caselli, F., Esquivel, G., Lefort, F. (1996). Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics, *Journal of Economic Growth*, 1 (3), 363-389.

Carhart, M.M., (1997). On Persistence in Mutual Fund Performance, *Journal of Finance*, 52/1, 57-82.

Chan, K.C., Yasushi, H., Lakonishok J., (1991). Fundamentals and Stock Returns in Japan. *Journal of Finance*, 46/5, 1739-1764.

Chan, K.C., Chen, N., Hsieh, D.A., (1985). An Exploratory Investigation of the Firm Size Level. *Journal of Financial Economics*, 14/3, 451-471.

Chen, F., Roll, R., Ross, S.A., (1986). Economic Forces and the Stock Market. *The Journal of Business*, 59/3, 383-403.

Cheng, B.S., Lai, T.W. (1997). An Investigation of Co-Integration and Causality Between Energy Consumption and Economic Activity in Taiwan. *Energy Economics* 19, 435-444.

Chow, G.C. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica* 28/3, 591-605.

Daniel, K., Sheridan, T., Wei, K.C.J. (2001). Explaining the Cross-Section of Stock Returns in Japan: Factors or Characteristics? *Journal of Finance*, 56/2, 743-766.

Dickey, D., Fuller, W. (1981). Likelihood Ratio Tests for Autoregressive Time Series with a Unit Root. *Econometrica* 49, 1057-1072.

Eichenbaum, E., Hansen, L., Singelton, K. (1988). A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice under Uncertainty. *Quarterly Journal of Economics*, 103, 57-78.

Engle, R.F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50, 987-1007.

Erol, U., Yu, E.S.H. (1987). On the Causal Relationship Between Energy and Income for Industrialized Countries. *Resources and Energy*, 11, 395-412.

Evans, M.K., Klein, L.R. (1967). *The Wharton Econometric Forecasting Model.* Philadelphia: Economics Research Unit, Department of Economics, Wharton School o Finance and Commerce, University of Pennsylvania.

Fabozzi, F.J., Francis, J.C., (1977). Stability Tests for Alphas and Betas over Bull and Bear Market Conditions. *Journal of Finance*, 32, 1093-1099.

Fama, F.E., French, K.R., (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, 33, 3-56.

Fama, F.E., French, K.R., (2004). The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspecitves*, 18/3, 25-46.

Fan, Y., Xu, Y.F. (2011). What has driven oil prices since 2000? A structural change perspective. *Energy Economics* 33 (6), 1082-1094.

Fry, R., Pagan, A. (2009). Sign Restrictions in Structural Vector Autoregressions: A Critical Review. NCER Working Paper 57.

Glasure, Y.U., Lee, A.R. (1997). Co-Integration, Error-Correction, and the Relationship Between GDP and Electricity: the Case of South Korea and Singapore. *Resource and Energy Economics*, 20, 17-25.

Godfrey, L.G. (1978). Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables. *Econometrica*, 46, 1293-1302.

Griffin, J. M., Ji, S., Martin, J.S., (2003). Momentum Investing and Business Cycle Risk: Evidence from Pole to Pole. *Journal of Finance*, 58, 2515–2547.

Hahn, J., Kuesteiner, G. (2002). Asymptotically Unbiased Inference for a Dynamic Panel Model with Fixed Effects when Both n and T Are Large. *Econometrica*, 70 (4), 1639-1657.

Hahn, J., Lee, H., (2006). Yield Spreads for Alternative Risk Factors for Size and Book-to-Market. *The Journal of Financial and Quantitative Analysis*, 41/2, 245-269.

Hamilton, J.D. (2003). What Is an Oil Shock? *Journal of Econometrics* 113(2), 363-398.

Hamilton, J.D. (2008). Understanding Crude Oil Prices. NBER Working Paper No. 14492.

Hamilton, J.D. (2009). Causes and consequences of the oil shock of 2007-2008.*Brookings Papers on Economic Activity*, pp. 215-261.

Hamilton, J.D., Herrera, A.M. (2004), Oil Shocks and Aggregate Macroeconomic Behavior: The Role of Monetary Policy. *Journal of Money, Credit, and Banking* 36, 265-286.

Hamilton, J.D., Wu, J.C. (2011). Risk Premia in Crude Oil Futures Prices. Working Paper. University of California at San Diego.

Hansen, B.E.. (2001). The New Econometrics of Structural Change: Dating Breaks in U.S. Labor Productivity. *Journal of Economic Perspectives*, 15, 117-128.

Hansen, L.P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50, 1029-1054.

Hauk Jr, W.R., Wacziarg, R. (2009). A Monte Carlo Study of Growth Regressions. *Journal of Economic Growth* 14, 103-147.

Hausman, J.A. (1978). Specification Tests in Econometrics, *Economtetrica*, 46, 1251-1271.

Hicks, J.R. (1937). Mr. Keynes and the Classics. *Econometrica*, 5(2), 147-159.

Hineline, D.R. (2008). Parameter Heterogeneity in Growth Regressions. *Economics Letters*, 101, 126-129.

Hotelling, H. (1931). The Economics of Exhaustible Resources. *The Journal of Political Economy*, 39 (2), 137-175.

Huang, B.N., Hwang, M.J., Yang, C.W. (2008). Causal Relationship Between Energy Consumption and GDP Growth Revisted: A Dynamic Panel Data Approach. *Ecological Economics*, 67, 41-54.

IEA (2006). World energy outlook 2006. International Energy Agency, OECD Publication Service, OECD, Paris. www.iea.org accessed 26.11.2010.

IEA (2009). World energy outlook 2009. International Energy Agency, OECD Publication Service, OECD, Paris. www.iea.org accessed 29.11.2010.

IEA (2010). Energy Poverty: How to Make Modern Energy Access Universal? Early Excerpt of the World Energy Outlook 2010. International Energy Agency, OECD Publication Service, OECD, Paris. www.iea.org accessed 26.11.2010.

Islam, N. (1995). Growth Empirics: A Panel Data Approach. *The Quarterly Journal of Economics*, 110 (4), 1127-1170.

Jagadeesh, N., Titman, S., (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance*, 48/1, 65-91.

Jarque, C.M., Bera, A.K. (1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review*, 55 (2), 163-172.

Jemelkova, B., Toman, M. (2003). Energy and Economic Development: An Assessment of the State of Knowledge. Discussion Paper 03-13, Resources for the Future, Washington D.C.

Judson, R.A., Owen, A.L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. *Economics Letters*, 95, 9-15.

Juvenal, L., Petrella, I. (2011). Speculation in the Oil Market. *Federal Reserve Bank of St. Louis Working Paper Series*, No. 12011-027E.

Kaufmann, R.K. (2011). The role of market fundamentals and speculation in recent price changes for crude oil. *Energy Policy* 39, 105-115.

Kelly, P.J., (2003). Real and Inflationary Macroeconomic Risk in the Fama and French Size and Book-to-Market Portfolios. Working Paper, W.P. Carey School of Business, Arizona State University.

Kilian, L (2009). Not all Oil Price Shocks are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market. *American Economic Review* 99:3, 1053-1069.

Kilian, L., Hicks, B. (2012). Did Unexpectedly Strong Economic Growth Cause the Oil Price Shock of 2003-2008? *forthcoming: Journal of Forecasting*.

Kilian, L., Murphy, D.P. (2011). The Role of Inventories and Speculative Trading in the Global Market for Crude Oil. Forthcoming, *Journal of the European Economic Association*.

Kiviet, J.F. (1995). On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models. *Journal of Econometrics*, 68, 53-78.

Kleiber, C., Zeilies, A. (2008). *Applied Econometrics With R*. Springer.

Kraft, J., Kraft, A. (1978). On the Relationship Between Energy and GNP. *Journal of Energy and Development*, 3, 401-403.

Krugman, P., 1998. Japan's Trap. Retrieved from *http://web.mit.edu/krugman/www/japtrap.html*

Krugman, P. (2008). The Oil Nonbubble. Retrieved from: *http://www.nytimes.com/2008/05/12/opinion/12krugman.html*

Kubota, K., Takehara, H., (1996). Cross-Section Risk and Return of Tokyo Stock Exchange Firms. *Advances in Pacific Basin Financial Markets*, 2, 273–305.

Kubota, K., Takehara, H. (1997). Common Risk Factors of Tokyo Stock Exchange Firms. *Advances in Pacific Basin Financial Markets*, 3, 273–302.

Kwiatkowski, D., Philipps, P.C.B; Schmidt, P.; Shin, Y. (1992). Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. *Journal of Econometrics* 54, 159-178.

Kydland, F.E., Prescott, E.C. (1982). Time to Build and Aggregate Fluctuations. *Econometrica*, 50(6), 1345-1371.

Kydland, F.E., Prescott, E.C. (1996). The Computational Experiment: An Econometric Tool. *Journal of Economic Perspectives*, 10(1), 69-85.

Lamla, M. J., Sarferaz, S. (2012). Updating Inflation Expectations. *KOF Working Paper No. 301.*

Lee, C.C. (2006). The Causality Relationship Between Energy Consumption and GDP in G-11 Countries Revisted. *Energy Policy*, 34, 1086-1093.

Lee, C.C., Chang, C.P. (2007). Energy Consumption and GDP Revisted: a Panel Analysis of developed and developing Countries. *Energy Economics*, 29, 1206-1223.

Lettau, M., Ludvigson S., (2001). Resurrecting the (C)CAPM: A Cross-Sectional Test when risk premia are Time-Varying. *Journal of Political Economy*, 109, 1238-1287.

Levy, R.A., (1974). Beta Coefficients as Predictors of Returns. *Financial Analysts Journal*, 30 (1), 61-69.

Liew, J., Vassalou, M., (2000). Can Book-to-Market, Size, and Momentum be Risk Factors that Predict Economic Growth? *Journal of Financial Economics*, 57, 221-245.

Lindenberger, D., Kummel, R. (2002). Energy-Dependent Production Functions and the Optimization Model "Prise" of Price-Induced Sectoral Evolution. *International Journal of Thermodynamics*, 5 (3), 101-107.

Liu, C., Lee, Y., (2001). Does the Momentum Strategy Work Universally? Evidence from the Japanese Stock Market. *Asia-Pacific Financial Markets*, 4, 321-339.

Lombardi, M.J., Van Robays, I. (2011). Do Financial Investors Destabilize the Oil Price? *ECB Working Paper Series*, No. 1346.

Lucas, R.E. Jr., (1973). Some International Evidence of Output-Inflation Tradeoffs. *American Economic Review*, 63 (3), 326-334.

Lucas, R.E. Jr., (1976). Econometric Policy Evaluation: A Critique. *Carnegie Rochester Conference Series on Public Policy*, 1, 19-46.

Lucas, R.E. Jr., (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, 22, 3-42.

Lütkepohl, H. (2004). Vector Autoregressive and Vector Error Correction Models. In: H. Lütkepohl, M. Krätzig (eds.), Applied Time Series Econometrics. Chapter 4, pp. 159-196. Cambridge University Press. Cambridge.

MacKinlay, A.C., Richardson, M.P., (1991). Using Generalized Method of Moments to Test Mean-Variance Efficiency. *Journal of Finance*, 46/2, 511-527.

Mankiw, N.G. (2006). The Macroeconomist as Scientist and Engineer. NBER Working Paper, 12349, Massachusetts.

Mankiw, N.G., Romer, D., Weil, D.N. (1992). A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics*, 107 (2), 407-437.

Masters, M. (2009). Testimony before the Commodity Futures Trading Commission. *Technical Report, Commodity Futures Trading Commission.*

Merton, R.C. (1973). An Intertemporal Asset Pricing Model, *Econometrica*, 41/5, 867-887.

Modigliani, F. (1944). Liquidity Preference and the Theory of Interest and Money. *Econometrica*, 12(1), 45-88.

Morana, C. (2013). Oil Price Dynamics, Macro-Finance Interactions and the Role of Financial Speculation. *Journal of Banking and Finance*, 37 (1), 206-226.

Nakaso, H., (2001). The Financial Crisis in Japan during the 1990s: How the Bank of Japan Responded and the Lessons Learnt. Bank of International Settlement Papers No. 6, BIS Basel.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49 (6), 1417-1426.

Oh, W., Lee, K. (2004). Energy Consumption and Economic Growth in Korea: Testing the Causality Relation. *Journal of Policy Modeling*, 26, 973-981.

Paul, S., Bhattacharya, R.N. (2004). Causality Between Energy Consumption and Economic Growth in India: a Note on Conflicting Results. *Energy Economics*, 26, 977-983.

Persson, T., Tabellini, J.B. (2008). The Growth Effect on Democracy: Is it Heterogenoeous and How Can it Be Estimated. Chap. 13 in *Institutions and Economic Performance*, ed. E. Helpman. Cambridge: Harvard University Press.

Pfaff, B. (2008). VAR, SVAR and SVEC Models: Implementation Within R Package vars. *Journal of Statistical Software*, 27 (4).

Phillips, K. L., Spencer, D. E. (2010): Bootstrapping Structural VARs: Avoiding a Potential Bias in Confidence Intervals for Impulse Response Functions. MPRA Working Paper.

Philipps, P.C.B, Perron, P. (1988). Trends and Random Walks in Macroeconomic Time Series. *Biometrika* 75, 335-346.

Rebelo, S. (1991). Long-Run Policy Analysis and Long-Run Growth. *Journal of Political Economy*, 99, 500-521.

R Development Core Team (2014). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. www.R-project.org accessed 8.1.2014.

Rodrik, D., Wacziarg, R. (2005). Do Democratic Transition Produce Bad Outcomes?, *American Economic Review*, 95(2), 50-55.

Romer, P.M. (1986). Increasing Returns and Long-Run Growth, *Journal of Political Economy*, 94 (5), 1002-1037.

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

Sargent, T.J., (1971). A Note on the Accelerationist Controversy. *Journal of Money Credit and Banking*, 3, 721-725.

Sargent, T.J., (1973). Rational Expectations, the Real Rate of Interest, and the Natural Rate of Unemployment. *Brookings Papers on Economic Activity*, 2, 429-472.

Sargent, T.J., (1976). A Classical Macroeconometric Model for the United States. *Journal of Political Economy*, 84, 207-237.

Schmidt, P.S., von Arx U., Schrimpf A., Wagner, A.F., Ziegler, A., (2011). On the Construction of Common Size, Value and Momentum Factors in International Stock Markets. Economics Working Paper Series 11/141, ETH Zurich.

Sims, C.A. (1980). Macroeconomics and Reality. *Econometrica* 48(1), 1-48.

Sims, C.A. (1986). Are Forecasting Models Usable for Policy Analysis? *Minneapolis Federal Reserve Bank Quarterly Review* ,10, 2-16.

Singleton, K.J. (2011). Investor Flows and the 2008 Boom/Bust in Oil Prices. Working paper. Stanford University.

Solow, R. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics*, 70 (1), 65-94.

Solow, R. (2008). The State of Macroeconomics. *Journal of Economic Perspectives*, 22 (1), 243-246.

Soytas, U., Sari, R. (2003). Energy Consumption and GDP: Causality Relationship in G-7 Countries and Emerging Markets. *Energy Economics*, 25, 33-37.

Stock, J.H., Watson, M.W., (1996). Evidence on Structural Instability in Maroeconomic Time Series Relations. *Journal of Business and Economic Statistics*, 14(1), 11-30.

Summers, L.H., (1991). The Scientific Illusion in Empirical Macroeconomics. *Scandinavian Journal of Economics*, 93 (2), 129-148.

Tang, K, Xiong, W. (2011). Index Investment and Financialization of Commodities. Working Paper. Princeton University.

Vassalou M. (2003). News Related to Future GDP Growth as a Risk Factor in Equity Returns. *Journal of Financial Economics*, 68, 47–73.

Vu Thang Pham Long, (2007). Constructing Fama-French Factors from style indexes: Japanese evidence. *Economics Bulletin* 7/7, 1-10.

Walid, E.M., (2009). New Evidence on Risk Factors, Characteristics and the Cross-Sectional Variation of Japanese Stock Returns. *Asia-Pacific Financial Markets*, 16, 33-50.

Walid, E. M., Ahlem, E.M., (2009). New Evidence on the Applicability of Fama and French Three-Factor Model to the Japanese Stock Market. Working paper, Osaka University.

Yu, S.H., Jin, J.C. (1992). Cointegation Tests of Energy Consumption, Income and Employment. *Resource and Energy*, 14, 186-190.

Yu, E.S.H., Choi, P.C.Y., Choi, J.Y. (1988). The Relationship Between Energy and Employment: a Re-examination. *Energy Systems Policy*, 11, 287-295.

Zeileis, A., Leisch, F., Kleiber, C., Hornik, K. (2005). Monitoring Structural Change in Dynamic Econometric Models. *Journal of Applied Econometrics*, 20, 99-121.