ETH zürich

KOF Dissertation Series

Essays on Business Cycles and Sectoral Dynamics

Anna Pauliina Sandqvist

Diss. ETH No. 24987

KOF Dissertation Series, No. 35, 2018

KOF Swiss Economic Institute

Imprint

Publisher

KOF Swiss Economic Institute, ETH Zurich

DISS. ETH NO. 24987

ESSAYS ON BUSINESS CYCLES AND SECTORAL DYNAMICS

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

ANNA PAULIINA SANDQVIST

Dipl. oec. in Economics, University of Hamburg

born on 01.07.1985

citizen of Finland

accepted on the recommendation of

Prof. Dr. Jan-Egbert Sturm, examiner

Dr. habil. Klaus Abberger, co-examiner

2018

Preface

This thesis became reality with support and guidance of many individuals. First, I would like to express my gratitude towards my supervisor Professor Jan-Egbert Sturm for giving me this opportunity and supporting and advising me during the whole time. During writing this thesis I was employed at KOF as a researcher and enrolled as doctoral student at ETH Zurich. The mixture of institute's work tasks and scientific research projects was an interesting and highly instructive combination. His constructive comments and suggestions kept me motivated and helped me to improve the quality of my projects. I am also deeply thankful to Dr. habil. Klaus Abberger, my immediate superior as a head of the group *Business tendency surveys* and co-supervisor of my thesis, for sharing his knowledge and his guidance. The fruitful discussions with him, his advice and his encouragement helped me to great extent to accomplish this work.

I also thank my ex-colleague and co-author Boriss Siliverstovs for his contribution to our research project and for his advice and help, in general. I am deeply grateful that he shared his widespread knowledge and his extensive experience with me.

It is a genuine pleasure to express my deep sense of thanks to all my friends and colleagues at work, who have also contributed to the preparation of this thesis. I acknowledge with thanks their support, being there for a talk and giving their advice and encouragement.

It is my privilege to thank my family, friends and especially Thomas for the constant support and faith in me.

Anna Pauliina Sandqvist

Contents

List of Figures

List of Tables

Summary of the thesis

The aim of this thesis is to gain new insights on business cycles, and particularly on the role of sectoral dynamics. Chapter 2 examines the time-varying characteristics of interindustry comovement and its determinants. Chapter 3 studies the meaning of common information for sectoral comovement. Chapter 4 analyses the effects of unusual weather on consumer spending and on the aggregate volatility of retail trade sector. Chapter 5 deals with an issue relating to the measurement of business cycles; more specifically, with the problem of identification of outliers in case of small samples and varying distributional characteristics.

Chapter 2 employs a novel, multivariate dynamic conditional correlation (DCC) general autoregressive conditional heteroskedasticity (GARCH) framework to study the dynamics of comovement across manufacturing sectors both in the United States and in Germany. This approach allows us to account not only for non-constant variances, but also for possible asymmetries in conditional volatilities and correlations. We find that comovement across sectors is not stable, but shows irregular movements. In particular, contractions tend to be more synchronized than expansions in the manufacturing sector. Furthermore, we examine the role of various aggregate factors in the fluctuations in conditional correlations. Our findings reveal that both the non-constant variability of common factors, as well as the changes in the effects of these factors, play a role in the fluctuations in sectoral comovement.

In Chapter 3 we address the meaning of common information for sectoral co-movement by applying new quantification of firms' expectations about productivity developments. We employ micro data from business tendency surveys in order to construct novel sector- and aggregate-level measures of anticipated productivity changes. We find that firms in different sectors do, in fact, base their expectations on similar information. Particularly, the common information component is found to explain a considerable share of the sectoral correlations. Additionally, our findings reveal that there is a great deal of heterogeneity in the reactions of disaggregated manufacturing sectors to changes in expectations. Especially, expectation shocks are found to propagate mainly through the sectors producing capital goods and consumer durables.

Chapter 4 analyses how weather anomalies cause fluctuations in consumer spending by means of a comprehensive periodic analysis at monthly frequency. We account for both contemporaneous as well as lagged effects and use precise weather measures. Moreover, we test for long-run restrictions and quantify the weather effects borrowing the approaches of Boldin and Wright (2015). In addition, we develop a stylized model, based on consumer choice, to illustrate how abnormal weather affects the utility derived from consumption and therefore, leading into intertemporal shifts in consumption. Using retail data for Switzerland, the empirical findings support our theoretical considerations suggesting that weather anomalies, especially unusual temperatures, cause sizable intertemporal shifts in consumer spending as the long-run restrictions cannot be rejected. In particular, our findings indicate that weather effects manifest primarily through the seasons change channel: weather conditions in line with the coming season boost the purchases early in the season.

Outliers and influential observations are a frequent concern in various types of statistics, data analysis and survey data. Even though a large number of techniques to identify outlying observations have already been proposed, methods to deal with outliers in skewed or heavy-tailed data are still scarce. Hubert and Vandervieren (2008) propose an adjusted boxplot for skewed data. However, their method is not adjusted for heavy-tailness and it does not perform well for small samples. Nevertheless, in practice we often need to deal with small to moderate sample sizes and varying distributional characteristics. In Chapter 5, we propose a modification of the adjusted boxplot of Hubert and Vandervieren (2008) that is less sensitive to both heavy tails and for small samples sizes. We conduct a simulation study to illustrate the superior performance of the modified procedure under various models and we provide real data examples.

Zusammenfassung der Dissertation

Das Ziel dieser Dissertation ist es, neue Erkenntnisse über Konjunkturzyklen und insbesondere über die Rolle der sektoralen Dynamik zu gewinnen. In Kapitel 2 werden zeitvariable Charakteristika der sektoralen Gleichbewegung und ihre Determinanten untersucht. Kapitel 3 erforscht die Bedeutung der gemeinsamen Information für die Gleichbewegung der Branchen. Kapitel 4 analysiert die Auswirkungen ungewöhnlicher Witterungsbedingungen auf Konsumausgaben und auf die Gesamtvolatilität des Detailhandels. Kapitel 5 beschäftigt sich mit einem Thema, das sich auf die Messung von Konjunkturzyklen bezieht, und zwar mit der Identifizierung von Ausreissern in kleinen Stichproben und in Datensätzen mit variierenden Verteilungsmerkmalen.

Kapitel 2 verwendet eine neuartige DCC-GARCH-Methode, um die Dynamik der Verbindungen zwischen den Sektoren des Verarbeitenden Gewerbes, sowohl in den Vereinigten Staaten als auch in Deutschland, zu untersuchen. Dieser Ansatz ermöglicht es uns, nicht nur die zeitvarianten Volatilitäten, sondern auch mögliche Asymmetrien in den bedingten Volatilitäten und Korrelationen zu berücksichtigen. Wir stellen fest, dass die Gleichbewegung der Sektoren nicht stabil ist, sondern unregelmässige Fluktuationen aufweist. Diese Analyse zeigt insbesondere auf, dass die Kontraktionen in der Regel synchronisierter sind als die Expansionen im Verarbeitenden Gewerbe. Darüber hinaus untersuchen wir die Rolle verschiedener Faktoren für die Fluktuationen in den bedingten Korrelationen. Unsere Ergebnisse zeigen, dass sowohl die nicht-konstante Variabilität der gemeinsamen Faktoren als auch die Veränderungen in den Auswirkungen dieser Faktoren für die Schwankungen in den sektoralen Gleichbewegungen von Bedeutung sind.

In Kapitel 3 befassen wir uns mit der Bedeutung der gemeinsamen Information für die sektoralen Gleichbewegungen mit Hilfe einer neuen Quantifizierungsmethode der Erwartungen der Unternehmen über die Entwicklung der Produktivität. Wir benutzen Mikrodaten der Konjunkturumfragen, um neuartige Masse für einzelne Subbranchen als auch für die aggregierte Branche zu konstruieren. Wir finden heraus, dass Firmen aus verschiedenen Sektoren ihre Erwartungen tatsächlich auf ähnliche Informationen stützen. Zudem kann die gemeinsame Komponente der Erwartungen einen beträchtlichen Teil der sektoralen Korrelationen erklären. Unsere weiteren Erkenntnisse zeigen, dass die Reaktionen auf veränderte Erwartungen zwischen Subbranchen des Verarbeitenden Gewerbes differenzieren. Insbesondere werden Erwartungs-schocks vor allem durch den Investitionsgütersektor und den Gebrauchsgütersektor weiterverbreitet.

In Kapitel 4 wird mittels einer umfassende periodische Analyse mit Monatsdaten untersucht, ob Wetteranomalien Schwankungen im Konsumverhalten verursachen. Wir berücksichtigen sowohl zeitgleiche als auch verzögerte Effekte und benutzen präzise Masse für das ungewöhnliche Wetter. Darüber hinaus benutzen wir die Ansätze von Boldin and Wright (2015), um Restriktionen, die die langfristen Effekte beschränken, zu testen und um die Wettereffekte quantifizieren zu können. Zudem entwickeln wir ein theoretisches Modell, welches veranschaulicht, wie abnormale Witterungsbedingungen den Konsumnutzen beeinflussen und somit zu intertemporalen Verschiebungen führen können. Unsere empirischen Erkenntnisse mit den Schweizer Detailhandelsdaten unterstützen die theoretischen Überlegungen, dass Wetteranomalien, insbesondere ungewöhnliche Temperaturen, intertemporale Verschiebungen der Konsumausgaben verursachen. Insbesondere deuten unsere Ergebnisse darauf hin, dass sich Wettereffekte vor allem durch den Kanal des Saisonwechsels offenbaren: Wetterbedingungen, die im Einklang mit der kommenden Saison sind, kurbeln die Einkäufe früh in der Saison an.

Ausreisser und einflussreiche Beobachtungen sind ein häufiges Problem bei verschiedenen Arten von Statistiken, Datenanalysen und Befragungsdaten. Auch wenn bereits eine grosse Anzahl von Techniken vorgeschlagen worden ist, um auffällige Beobachtungen zu identifizieren, sind Methoden zur Erkennung von Ausreissern in asymmetrisch verteilten Daten oder in Datensätzen mit starken Ränder nach wie vor rar. Hubert und Vandervieren (2008) schlagen einen angepassten Boxplot für schiefe Datensätze vor. Ihre Methode ist jedoch nicht optimal für Stichproben mit breiten Rändern oder für kleine Datensets. In der Praxis müssen wir jedoch oft mit kleinen bis mittleren Stichprobengrössen und variierenden Verteilungscharakteristika umgehen. In Kapitel 5 schlagen wir eine Modifikation des angepassten Boxplots von Hubert und Vandervieren (2008), welche weniger empfindlich auf breite Ränder und kleine Stichproben reagiert. Wir führen eine Simulationsstudie durch, um die überlegene Leistungsfähigkeit des modifizierten Verfahrens bei verschiedenen Modellen zu veranschaulichen und präsentieren zudem reale Datenbeispiele.

Chapter 1

Introduction

Understanding and explaining business cycles is one of the most challenging topics in macroeconomics. Research on these fluctuations began on a large scale during the middle of the previous century at the National Bureau of Economic Research (NBER) by Burns and Mitchell in 1964. Since then, analysing business cycles has become a crucial topic in macroeconomics. Throughout the years, numerous hypotheses and explanations for these fluctuations have been introduced. Nevertheless, there are still many open issues relating to how business cycles evolve.

The term *business cycle* refers to, on one hand, joint movement of aggregate variables such as output, investment and consumption and on the other hand, to synchronized movements of disaggregated sectoral variables. Therefore not only *aggregate comovement*, but also *sectoral comovement* is a main feature of business cycles. However, literature on economic fluctuations has so far chiefly concentrated on examining comovement between aggregate variables. This is likely the case as standard business cycle models assume only one economic sector. Obviously, within such a framework, industrial comovement cannot be addressed. Additionally, within two-sector models (with investment and consumption good sectors), generating sectoral comovement has turned out to be difficult, even in the case of aggregate technology shocks as discussed by Christiano and Fitzgerald (1998), for instance. This thesis aims to reveal new insights about business cycles by elaborating the role of sectoral dynamics.

Sectoral comovement is also a crucial determinant of aggregate volatility: the variability of aggregate output depends not only on the volatility of individual sectors but also, to a great extent, on correlations between them. As Foerster et al. (2011) point out, high volatility of the aggregate industrial production index is quite surprising, as it is constructed as a weighted average of numerous sectors, and the variability of sectors do not seem to be averaged out. This is owed to the fact that sectoral movements tend to be positively correlated and this matters a great deal. Shea (2002) demonstrates that comovement of sectors accounts for more than 85 percent of the variance of aggregated US manufacturing gross output. Foerster et al. (2011) also find the covariance terms to account for 55% to 70 % of the variation of aggregate US industrial production growth by applying quarterly data. Therefore, interindustry correlations are the key to explain aggregate volatility.

Furthermore, sectoral comovement helps to reveal the origins of economic fluctuations, and especially the relative importance of aggregate versus sectoral shocks. The joint behaviour of sectors is commonly interpreted as evidence for aggregate shocks causing the business cycle (e.g., Lucas (1977)), such as monetary policy or technology shocks affecting all industries in a similar manner. Yet the synchronized movement across sectors does not necessarily need to be a result of common shocks. Sector-specific shocks could also be propagated through sectors causing industries to move together. Long and Plosser (1983) showed that in a multisector framework with intersectoral linkages (input-output dependencies), independent sectoral shocks alone may induce correlation between sectors. However, as they assume a one-period delivery lag, the lagged correlations turn out to be higher than the contemporaneous ones. Moreover, they assume full depreciation of capital and when this assumption is relaxed, many properties no longer generalize as noted by Rebelo (2005). Similarly Horvath (1998, 2000), Dupor (1999), Shea (2002), Carvalho (2008) and Foerster et al. (2011) examine the role of intersectoral dependencies transmitting sector-specific (uncorrelated) shocks within structural multisector models. Nonetheless, even if the rather unrealistic assumption[1] of contemporaneous sectoral linkages

[1] Unrealistic because it assumes hat goods are produced in period t in one sector and being used as input or capital at the same period in a another sector.

is placed, they tend not to be strong enough, especially at the more detailed level of disaggregation, in order to account for a realistic level of interindustry comovement (see, Foerster et al. (2011)). Further mechanisms such as inventories (Cooper and Haltiwanger (1990)) and trade credit (Raddatz (2010)) have been proposed. They play though at best only a moderate role in explaining synchronized sectoral movements.

To understand the synchronized behaviour across industries, the first issues to be elucidated include the level of sectoral comovement and, in particular, the manner in which it evolves over time. Although the level of interindustry comovement has already been documented quite comprehensively (see, for example, Long and Plosser (1987), Christiano and Fitzgerald (1998), Hornstein (2000), Cassou and Vázquez (2014)), little is known about its dynamic aspects as these papers do not address the level of comovement over time. Nevertheless, there is no reason to expect that these correlations across sectors could not change over time. Though only a few empirical works have addressed this issue so far simply by different subsamples (see, Romer (1991), Foerster et al. (2011)). This method offers, quite obviously, only a limited picture about the changes in correlations over time and it depends, to a great extent, on the definitions of the subsamples. More sophisticated approaches have hardly been applied. This is likely to be the case as measuring time-varying correlations is not clear-cut. It is of particular importance to account for the non-constant volatility when modelling conditional correlations, as the estimated correlations tend to be higher in times of increased volatility. Only by accounting for time-variant volatility can one distinguish whether higher correlations are due to higher variance or stronger covariance.

In Chapter 2 we explore the measurement and dynamics of sectoral business cycle comovement in more detail. To study the time-varying aspects accurately, we employ a novel, state-of-the-art dynamic conditional correlation (DCC) general autoregressive conditional heteroskedasticity (GARCH) framework. This approach allows us not only to account for non-constant variances, but also to account for possible asymmetries in conditional volatilities and correlations. We find that comovement across sectors is not stable, but shows irregular movements. Specifically, contractions tend to be more synchronized than expansions in manufacturing sector. Moreover, we examine the role of various aggregate factors for the fluctuations in conditional correlations. Our findings reveal that both the non-constant variability of common factors, as well as the changes in the effects of these factors play a role in the fluctuations in sectoral comovement.

A further proposition how the joint behaviour of industries might evolve is based on information complementarities or common information. Christiano and Fitzgerald (1998) discussed *information externalities* as a potential source for industrial correlations. However, Veldkamp and Wolfers (2007) were the first to offer a theoretical model for the role of common expectations, or aggregate information, for sectoral comovement. They argue that complementarity in information acquisition could also explain the so-called *excess comovement puzzle* first documented by Hornstein (2000) - an important, but often ignored, empirical finding - that production is stronger correlated across sectors than productivity. This stylized fact is clearly not in line with an aggregate technology shock being the main source of economic fluctuations. The explanation of Veldkamp and Wolfers (2007) is based on the idea that firms collect mostly aggregate information about productivity developments, as it is cheaper to acquire than sector-specific information. Based on similar anticipations, firms tend to make similar input and output decisions even though actual sectoral productivity developments differ. Nevertheless, there is hardly any empirical support for their hypothesis. This is most likely because data on expectations, especially at the sector level, is scarce in general.

In Chapter 3 we aim to fill this gap applying new quantification of firms' expectations about productivity changes. We employ micro data from business tendency surveys to construct novel sector- and aggregate-level measures of anticipated productivity changes. We find that firms in different sectors do, in fact, base their expectations on similar information. Particularly, the common information component is found to explain a considerable share of the sectoral output correlations. Our additional findings reveal that there is a great deal of heterogeneity in the reactions of disaggregated manufacturing sectors to changes in expectations. Especially, expectation shocks are found to propagate mainly through the sectors producing capital goods and consumer durables.

Sectoral variability also reveals information about the drivers behind the industry movements. The differences in volatility may be due to greater sectoral shocks, or stronger responses of these sectors to common shocks, or both. Lucas (1977) states that the production of durable goods, that is, consumer durables and capital goods, has more amplitude than the non-durable goods (intermediate goods and consumer non-durables). Furthermore, Mankiw (1985) assesses the role of the higher variability of consumer expenditures on durable goods for business cycles, and their sensitivity to real interest rates. Consumer spending on some non-durable consumption goods is also known to be quite volatile such as apparel, minor sports equipment and other fashion or seasonal products. In general, the higher volatility of durable consumer goods and fashion products is because household demand for nonessential goods fluctuates considerably whereas a sizeable share of household consumption is related to necessities such as food, housing, health, transportation and education services. These necessities are documented to have a rather constant growth over time. Thus retail sales, and especially non-food categories, exhibit much higher volatility than total consumption expenditures. One factor for this higher variability, which is also regularly discussed in retail business and business press, is unusual weather. Exceptional weather conditions are argued to have an impact on consumer decisions and business activity, being thus one of the main causes for the transitory shifts.

Understanding how unusual weather affects consumer spending is of importance for several reasons. On one hand, it reveals how consumer decisions are affected by abnormal weather conditions through various channels. Linden (1962) already noted that unusual weather conditions cause shifts in the timing of purchases, they generate purchases that might otherwise not occur, or they cause a permanent loss of demand. However, these effects could be (partly) caught up in the following month(s). On the other hand, it is a well-known fact that consumer expenditure has a substantial transitory component which cannot be explained by changes in fundamental economic factors, such as income or interest rates. Unusual weather is assumed to be one of the main reasons for it. The transitory component can be problematic, as it can distort the measurement of the effects of fundamental factors on spending. For instance, the estimation of elasticity of intertemporal substitution (EIS), one of the central parameters of macroeconomic and finance models, is affected by this issue. Furthermore, as retail sales is one of the main economic indicators, its movements are followed closely by central banks and economic analysts. It is crucial to discern whether or not the observed changes are rather transitory (and possibly followed by a rebounce in following month) or if they reflect genuine changes in underlying factors. Therefore, impacts caused by exceptional weather are relevant for business cycle analysis and monitoring current economic conditions, as well as for making projections in the future.

In Chapter 4 we explore the effects of unusual weather on consumer spending. We contribute to the limited literature by a comprehensive periodic analysis at monthly frequency, accounting for both contemporaneous and lagged effects and by using precise weather measures. Moreover, the application of long-run restrictions and the quantification of these effects, borrowing the approaches of Boldin and Wright (2015) who examine the weather-adjustment of employment data, is new in weather-related consumer spending literature. In addition, we develop a stylized model based on consumer choice in order to illustrate how abnormal weather affects the utility which, therefore, results in intertemporal shifts in consumption. Using retail data for Switzerland, the empirical findings support our theoretical considerations which suggest that weather anomalies (especially unusual temperatures) cause sizeable shifts in consumer spending and account for a considerable share of the variability of retail sales. Our findings also indicate that weather effects manifest mainly through the seasons change channel: exceptionally warm temperatures in early spring as well as unusually cold conditions in late summer and early autumn are generally associated with higher sales than usual.

Moreover, the measurement of business cycles is a nontrivial issue. Outliers and influential observations are a frequent concern in several types of statistics, data analysis and survey data. Owing to the fact that most of the outlier detection methods assume that the underlying population is normally distributed, in the case of non-normal distribution too many or the wrong observations are potentially declared as outlying. Thus, if the base distribution itself is assumed to be asymmetric or heavy-tailed, this needs to be taken into account in outlier analysis. For example, resistant rules, also known as boxplot, which was first proposed by Tukey (1977), are very popular in outlier labelling. The problem with skewed data is, however, that one tends to identify (too) many outliers in the long tail but hardly any in the short tail. Yet in the case of a symmetric, but heavy-tailed sample, numerous observations tend to be outside the fences on both sides. The same issues also concern other outlier detection techniques such as relative distance based methods.

Even though large number of techniques to identify outlying observations have already been proposed, methods to deal with outliers in skewed or heavy-tailed data are still scarce. Hubert and Vandervieren (2008) propose an adjusted boxplot for skewed data. However, their method is not adjusted for heavy-tailness and it does not perform well for small samples. Nevertheless, in practice we often need to deal with small to moderate sample sizes and varying distributional characteristics. Chapter 5 addresses outlier detection in case of small samples from skewed and heavy-tailed distributions. We propose a modification of the adjusted boxplot of Hubert and Vandervieren (2008) that is less sensitive to heavy tails and to small samples sizes. With our modification, we widen the practical application of the adjusted boxplot method. We conduct a simulation study to illustrate the superior performance of the modified procedure under various models and present real data examples.

Chapter 2

Dynamics of sectoral business cycle comovement[1]

[1]This chapter is based on Sandqvist (2017)

2.1 Introduction

Sectoral or interindustry comovement is a key feature of business cycles. A high level of sectoral comovement is commonly interpreted as evidence for strong direct impacts of aggregate shocks to all sectors as well as pass-through of common and sectoral shocks to whole economy due to sectoral linkages. Furthermore, the volatility of aggregate output depends not only on the volatility of individual sectors but also to a great extent of correlations between them. Shea (2002) shows that comovement between sectors accounts for more than 85 percent of the variance of aggregated US manufacturing gross output. Also Foerster et al. (2011) find the covariance terms to be crucial for the variation of aggregate US industrial production growth.

Although interindustry comovement has already been documented quite comprehensively (see, for example, Long and Plosser (1987), Christiano and Fitzgerald (1998), Hornstein (2000), Cassou and Vázquez (2014)), little is known about its dynamic aspects as these papers do not study patterns of comovement over the years nor depending on the state of the business cycle. Yet, there are reasons to expect that properties and dynamics of comovement might change over time in general as well as differ across the business cycle. First, there could be positive time trend as sectors could become more correlated over time due to stronger interindustry linkages in the course of years. On the other hand, there could be also a negative time trend as a result of lower magnitude of common factors. Foerster et al. (2011) document a decline in average pairwise correlations of sectoral growth rates in the US in the 1984-2007 period compared to the pre-1984 years. They find that the lower comovement is due to lower variance of common shocks what, in turn, has lead to decreased variance of the growth of industrial production in the same period. Second, the level of comovement might fluctuate. This could be because of asymmetric reactions to unfavourable and favourable shocks or differing reactions regarding the magnitude of a shock. Third, if the sensibility of industries to common shocks varies over time, this would also cause movements in sectoral correlations. For instance, if common factors become more important compared to sectoral shocks, this would results in higher comovement.

Yet, measuring time-varying correlations is not clear-cut as they cannot be observed but need to be estimated. Furthermore, it is necessary to account for the possibly non-constant volatility when modelling conditional correlations as the estimated correlations tend to be higher in times of increased volatility. Only by accounting for time-variant volatility one can distinguish whether higher correlations are due to higher volatility or stronger comovement of variables.

In this study we employ a multivariate dynamic conditional correlation (DCC) general autoregressive conditional heteroskedasticity (GARCH) framework, introduced by Engle and Sheppard (2001) and Engle (2002), to study the time-varying correlations of industrial production growth rates between manufacturing sectors in the US and Germany. The advantage of DCC-GARCH is that we can examine possible changes in conditional correlations depending on state or circumstances of the economy as well as generally over time. Furthermore, we can account for asymmetries in the conditional volatilities as well as also in the conditional correlations i.e. allow negative and positive shocks to have different effects. This is important because if an asymmetric process is modelled by the standard symmetric model, the estimated conditional variance respectively correlation after an unfavourable shock would be underestimated whereas the conditional volatility respectively correlation would be too high after a favourable shock.

Our results reveal that variance of industrial production growth in manufacturing sectors in the US tend to be asymmetric i.e. most industries react stronger to unfavourable than favourable shocks. However, we find only limited evidence for asymmetries in correlations what implies that most sector pairs react similarly to common shocks, irrespective of the sign of shocks. Furthermore, we show that the sectoral comovement in the US and German manufacturing is not constant but shows irregular movements. Interestingly, the correlations seem to have increased considerably during some economic downturns, especially during the recession in 2007-09, but there are also recessionary periods in which the comovement hardly changes.

Given that sectors comove because of the direct effects of common factors and pass-through of (aggregate and sectoral) shocks, the changes over time in interindustry comovement indicate altering magnitude of common shocks, changing sensitivity of sectors to common factors and/or varying strength of propagation of shocks. Yet, since the intensity of the pass-through depends mainly on sectoral linkages which are assumed to be rather constant or changing slowly, the first two reasons should be the main drivers of the fluctuations in interindustry correlations.

We examine first which aggregate factors are relevant for sectoral comovement in a regression analysis. Our findings for the US indicate that stock market index as well as business confidence are related to sectoral comovement. Especially, in times of pessimism the sectoral production decisions are more similar than in periods of high expectations. Also stock market volatility is found to play a role implying higher correlations in times of high volatility. Yet, for Germany, only the stock market volatility is found to have a significant impact on sectoral correlations. Furthermore, our findings suggest that time-varying variance of common factors as well as the altering impact of these factors is relevant for the fluctuations in sectoral comovement.

The remainder of the paper is organized as follows. First, in section 2 a literature overview is given. Section 3 presents the DCC-GARCH framework. In section 4 the research methodology and data are described and in section 5 the results of DCC-GARCH approach are presented. In section 6 we address the factors explaining the irregular movements in conditional correlations. Finally, section 7 offers conclusions.

2.2 Literature Overview

One of the pioneer works on sectoral comovement is the paper of Long and Plosser (1987). They study cross-industry comovement in the monthly US industrial production and apply factor analysis to examine the importance of aggregate versus sectoral shocks. They document the level of comovement by calculating average pairwise correlations and find the level of comovement to range between 0.07 and 0.28 for seasonally adjusted growth rates. Furthermore, they consider one- and two-factor models and find the explanatory power of the common shocks for each sector to be significant but rather weak. Foerster et al. (2011) study also the role of sectoral and aggregate shocks for industrial production in the US. They find the average pairwise correlation of sectoral monthly growth rates to be 0.19 for the whole sample period of 1972 to 2007. Moreover, they point out that the covariance terms are mainly responsible for the variation in the growth rate of aggregate industrial production index, whereas the variance terms of each sector play only minor role.

Christiano and Fitzgerald (1998) study the comovement in quarterly hours worked in the US two-digit industries. They estimate the level of correlation to be on average 0.55 among all industries. Furthermore, they show that the comovement of hours worked with total hours worked is higher across the durable goods manufacturing sectors (0.82) than across the nondurable manufacturing sectors (0.46)[1]. Also service sectors are comoving less with the general business cycle than the durable manufacturing but more than the nondurable goods manufacturing. However, they apply the reference series methodology, i.e. they calculate the correlation between a sectoral variable and its aggregate counterpart (which equals the sum of the sectoral series). Yet, the problem when using a reference series is that as the sectoral variable is a subaggregate of the reference series, they are per definition through the aggregation to some extent correlated with each other and therefore, one cannot distinguish to which extent to comovement is due to 'real' comovement and due to aggregation issue i.e. the measures of comovement based on reference series are more or less biased. Furthermore, Christiano and Fitzgerald (1998) analyse also some possible explanations for comovement without any definite findings. Hornstein (2000) documents also sectoral comovement in the US in the yearly data using both reference series as well as direct cross-industry measures with basic correlation measures. He considers besides the contemporaneous correlations also once-lagged respectively once-lead correlations. In most cases, however, the contemporaneous correlations tend to be the highest, what is not surprising as they use yearly data.

[1] Usually, following sectors are considered to be durable goods sectors: Wood, Metal, Machinery, Electronics, GlassStone, Electricals, Vehicles, Transportation, Furniture, Other. On the other hand, non-durable goods sectors include FoodTobacco, Textile, Apparel, Leather, Paper, Print, Petroleum, Chemicals, Rubber

To investigate changes in correlations over time various methodologies have been applied. A rather simple approach is the rolling window methodology, see for example Inklaar et al. (2008) and Papageorgiou et al. (2010). However, this approach is sensitive to the choice of the window length and cannot be applied to pinpoint (exact) time points when correlations change.

Yet, the most approaches to model time-varying variances, covariances and correlations originate from the finance literature. The generalized autoregressive conditional heteroskedasticity (GARCH) models were developed to model and estimate processes with non-constant variance and their multivariate extensions can also be applied to study conditional covariances and correlations. One of the first multivariate GARCH models was proposed by Bollerslev et al. (1988) as a direct generalizations of univariate GARCH model, the VEC model. As this model is heavily parameterized and demanding to estimate, further parametrizations for the time-varying covariance matrix have been introduced, such as the BEKK model of Engle and Kroner (1995). Like the VEC model also the BEKK model aims to directly estimate the conditional covariance matrix. However, as the number of parameters to be estimated in the BEKK specification is still high and the interpretation of parameters is not straightforward, empirical applications are rather rare and restricted to small dimensional systems.

Another class of multivariate GARCH models is based on nonlinear combinations of univariate GARCH models: constant conditional correlation (CCC) model of Bollerslev (1990) and the dynamic conditional correlation (DCC) model of Engle and Sheppard (2001) and Engle (2002). A very similar model was also introduced by Tse and Tsui (2002). These models overcome the problem of high parametrization and estimation issues at the cost of simpler model structure i.e. interdependencies between variances and among correlations are left out. Yet, they are very flexible considering the univariate variance modelling and offer intuitive interpretation of parameters. The assumption of standard scalar DCC-GARCH model tend to be, however, too restrictive for systems with a large set of variables. This is because the dynamics in the correlation process are forced to be the same for all variable pairs. Therefore, various generalizations of DCC has been proposed, for instance by

Billio et al. (2006) (Flexible DCC) as well as Cappiello et al. (2006) (Asymmetric Generalized DCC) where the scalar parameters are replaced by parameter matrices. But also with these models the number of parameters increase quickly with the size of the system and hence, various more restrictive specifications has been proposed (diagonal and symmetric, for example). All together, the problem with estimating conditional covariances respectively correlations for a large system is that the direct models (VEC or BEKK) are too complicated to be estimated while the scalar DCC is too restrictive, i.e. the assumption about the common dynamics is usually unrealistic.

The flexible DCC relaxes this assumption by restricting groups of variables to follow common dynamics. However, these groups needs to be defined beforehand. The diagonal version of the generalized DCC also allows for richer dynamics as for each variable a separate parameters are estimated and the correlation process of a pair of variables depends on the product of their parameters. However, one of the computational advantages of the DCC models, namely the possibility to apply correlation targeting which implies replacing the intercept matrix with a consistent estimate, is lost here.

The *MacGyver* method to deal with the high-dimensionality problem was proposed by Engle (2009). This appoarch is based on the estimates of the bivariate models which are aggregated through simple procedures, like mean or median. Another way to deal with large dimensions is the method of composite likelihood introduced for this context by Engle et al. (2007). This approach is very flexible but inefficient as it based on partial likelihood. A further alternative to a large system is to apply the Dynamic Equicorrelation (DECO) model of Engle and Kelly (2012). Yet, the assumption that all variable pairs have the same correlation on a given time point, as the correlation process is modelled between average correlation of the variable pairs, is quite restrictive.

Jondeau and Rockinger (2006) follow somewhat different approach and propose copula-based GARCH models to model conditional dependencies. Compared to the other multivariate GARCH models which usually assume data to follow multivariate normal or Student t-distribution, this copula-based approach can also take into account the possible skewness of data series. Furthermore, tail dependency could also be incorporated through the copula.

A related approach to these multivariate GARCH models is the panel data model with conditional heteroskedasticity and cross-sectional dependence of Cermeno and Grier (2008). Unlike in DCC-GARCH, the conditional mean (as well as conditional variance) model has to be the same for each series. Furthermore, the dynamics in conditional variance as well as in conditional covariance are forced to be the same (this keeps the number of parameters to be estimated small).

The empirical applications of multivariate GARCH models to comovement analysis are still scarce. Ho et al. (2009b) apply multivariate GARCH to study the asymmetric volatility and time-varying correlations between sectors of US industrial production. They find that negative shocks have greater impact on future volatilities. Furthermore, they find that evidence for time-varying conditional correlations. In Ho et al. (2009a) also a multivariate asymmetric GARCH approach is used to analyse volatility dynamics in the UK business cycle. They find evidence that conditional volatilities as well as correlations tend to be higher during UK recession periods.

2.3 Econometric methodology

During the last decade the DCC-GARCH framework, which was introduced by Engle and Sheppard (2001) and Engle (2002), based on the earlier works of Engle (1982), Bollerslev (1986) and Bollerslev (1990), has become more popular. The advantage of DCC-GARCH framework is that it has the flexibility of univariate GARCH but is not as complex as conventional multivariate GARCH models. The DCC-GARCH framework accounts for the time-varying volatility of the series and the conditional correlation matrix is time-dependent and therefore, it can be applied to study the dynamics in correlations. It has been applied to investigate, besides the vast finance related literature, dynamic comovement between stock market returns and policy uncertainty (Antonakakis et al., 2013) as well as output and prices (Lee, 2006). The advantage of DCC-GARCH is that we can examine possible changes in conditional correlations over time as well as depending on state or circumstances of the economy. Furthermore, with this approach we are also able study the asymmetric reactions of correlations to negative respectively positive shocks.

Let u_t be a data series with mean zero or residuals from a filtered time series

$$u_t \sim N(0, H_t).$$

For easier interpretation of the DCC-GARCH framework, H_t can be rewritten as:

$$H_t = D_t R_t D_t \tag{2.1}$$

where $D_t = diag\left\{\sqrt{h_{i,t}}\right\}$ is a diagonal matrix of time-varying standard deviations and R_t is a correlation matrix comprising conditional correlation coefficients. The standard deviations in matrix D_t are typically modelled to follow a univariate GARCH(1,1) (Bollerslev, 1986) of:

$$h_{i,t} = \omega_i + \alpha_i u_{i,t-1}^2 + \beta_i h_{i,t-1} \tag{2.2}$$

However, the model does not necessarily need to be the symmetric GARCH model but also other GARCH models can be incorporated. In this paper, we consider also the following models: exponential GARCH (Nelson, 1991), Asymmetric power ARCH (Ding et al., 1993) and The Glosten-Jagannathan-Runkle GARCH (Glosten et al., 1993). Furthermore, each series can have their own individual GARCH model as it is not necessary to model all with the same process. The standardized residuals are defined as $\varepsilon_t = u_t/\sqrt{h_t}$.

The correlation process in standard DCC model of Engle (2002) is given by

$$Q_t = (1 - a - b)S + a\varepsilon_{t-1}\varepsilon_{t-1}' + bQ_{t-1} \tag{2.3}$$

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1} \tag{2.4}$$

where $S = E[\varepsilon_t \varepsilon_t']$ is the unconditional covariance matrix of the standardized residuals, Q_t^* is a diagonal matrix with the square root of the diagonal elements of Q_t and

a and b are scalars. The parameter a of the equation (3) captures the effect of news on the correlation process whereas b is the smoothing parameter. The closer the sum of a and b is to unity, the stronger the persistence of the time-varying correlations. This is the *mean-reverting* DCC(1,1) model as long as $a + b < 1$.

In the estimation of the mean-reverting DCC model, *correlation targeting* is being applied i.e. the intercept matrix Ω is replaced by $(1 - a - b)S$. However, as Engle (2009) and Aielli (2013) note, this is only an approximation as S is not exactly $\bar{Q}$.

The estimation procedure of this system has three stages. First, univariate GARCH processes are estimated for each series. Second, the sample correlation matrix of the standardized residuals is calculated and third, the parameters of the correlation process are estimated.

An interesting extension of this model is the *asymmetric* DCC (aDCC) of Cappiello et al. (2006). In this model reactions of correlations to negative shocks can be greater than to positive shocks. Here the correlation process can be written as

$$Q_t = (1 - a - b)S + a\varepsilon_{t-1}\varepsilon'_{t-1} + bQ_{t-1} + \gamma\eta_{t-1}\eta'_{t-1} \tag{2.5}$$

where $\eta_t = min[\varepsilon_t, 0]$. The γ parameter measures the asymmetric effect. A positive value of γ indicates that the correlations increase more in response to negative than positive shocks.

2.4 Data and research methodology

In this paper, we use seasonally adjusted monthly industrial production index (IPI) data for US manufacturing on NAICS 3-digit level from Federal Reserve Economic Data base. The sample period runs from January 1972 to April 2015. The industrial production data for German manufacturing sectors (NACE 2-digit level) from Eurostat covers the period from January 1991 to April 2015. Even though manufacturing sector accounts for about 25 % of the total value added, it is responsible for most cyclical movements. Therefore, it is well suited for sectoral comovement analysis. The data contains 21 subsectors for each country (for exact definitions, see Tables A.1 and A.2 in the A Appendix). As most of these series are trending, we use transformation to first order log differences.

As we expect that the growth rates of industrial production index in the manufacturing subsectors are driven by different processes, we model each series by autoregressive moving average (ARMA) process chosen by the Bayesian Information Criterion (BIC)[2]. The diagnostic tests for the ARMA(l,k)-residuals are shown in Tables A.3 and A.4 in the A Appendix. The first two columns display the weighted Ljung-Box test (for a lag order of 20) for serial correlation of the residuals and mean-adjusted squared residuals. In most cases, serial correlation is still present in the squares of the residuals, i.e. in the variance, which indicates presence of conditional heteroskedasticity. The third column reports the results of a Lagrange multiplier (LM) test for ARCH with 10 lags and they also support the evidence of conditional heteroskedasticity for most of the residual series. Columns 4 to 6 show the results of normality tests. The null hypothesis of the Jarque-Bera test can be rejected at 5% significance level for all but one series indicating that most of the series are non-normal.

Given that the residual series seem to be heteroscedastic, a framework accounting for non-constant volatility is necessary for the analysis. As it is possible that the volatility dynamics are nonlinear, i.e. volatility reacts differently to negative and positive shocks, also asymmetric GARCH-models are considered: exponential GARCH (eGARCH), Asymmetric power ARCH (apARCH) and The Glosten-Jagannathan-Runkle GARCH (GJR-GARCH). If an asymmetric process is modelled by the standard symmetric model, the estimated conditional variance after a negative shock would be underestimated whereas the conditional volatility would be too high after a positive shock. The optimal GARCH(p,q)-order, model and distribution is chosen by BIC, however adjusted manually in some cases if the chosen model did not fit properly. Considered residual distribution assumptions include normal, skewed normal, t-student and skewed t-student distribution. The estimated coefficients for ARMA-GARCH models are presented in Tables A.5 and A.6 in the

[2]In some cases manual changes were taken if the chosen order did not seem to fit well, i.e. the coefficient of the highest order was estimated to be insignificant

A Appendix. For the US data, the asymmetric models are chosen more than half of the series. This indicates that variance of industrial production growth in manufacturing sectors in the US tend to be asymmetric i.e. most industries react stronger to unfavourable than favourable shocks (as the estimate for the asymmetry parameter is in all cases positive). In German data, however, there does not seem to be much asymmetry in the volatilities, as the symmetric GARCH is in most cases the most appropriate choice.

2.5 Empirical results

Since the correlation process is modelled as a scalar process in the basic DCC framework, only a single news impact parameter a and a single smoothing parameter b are being estimated. Hence, estimating one big model with many series would mean that the dynamics in correlations would be the same for all series. As we have 21 sectors for each country, we think this is too restrictive assumption. To overcome the problems with high-dimension models and stay flexible in modelling, we follow the MacGyver approach of Engle (2009) and estimate bivariate DCC(1,1)-(G)ARCH models with t-student distributed errors for all combinations of the series. Furthermore, we also account for the possible asymmetries in the conditional correlation and choose the asymmetric DCC(1,1) if the log likelihood is significantly higher than the one of the symmetric model. With 21 sectors, there are $21 \times 20/2 = 210$ bivariate models and thus, 210 series of dynamic correlations for each country.

2.5.1 Results for the US

The descriptive statistics of the estimated parameters of the bivariate DCC models are presented in Table 2.1. The results are satisfactory as the estimates for a are always between 0 and 0.188 and b estimates are also non-negative and smaller than unity. For some sector pairs the a estimate is relatively large and thus, the correlation process is quite erratic. This could indicate that these sector-pairs are rather weakly integrated. Further, the median for a is 0.016 and for b 0.888 and thus, the sum equals 0.904. This implies a rather moderate level of persistence in the correlation process.

The asymmetric version of the DCC model is preferred in 21 cases by the log-likelihood criterion. Given that the coefficient of γ is estimated to be positive, it indicates a higher reaction of the conditional correlations to negative than positive news for these sector pairs.

Table 2.1: Descriptive statistics of the all DCC model estimates for US data

Statistic	N	Mean	St. Dev.	Min	Median	Max
a	210	0.032	0.038	0.000	0.016	0.188
b	210	0.726	0.314	0.000	0.888	0.993
γ	21	0.207	0.100	0.068	0.210	0.440

In order to examine the interindustry comovement of the manufacturing sectors altogether, the mean of the pairwise conditional correlations is calculated. In Figure 2.1, the average conditional correlation across all the US manufacturing sectors is plotted with the 95 % confidence bands and the shaded areas indicate US recession as defined by the National Bureau of Economic Research. First of all, the mean of the average dynamic correlation over the sample period is moderate and around 0.187. This is in line with the previous literature given that Long and Plosser (1987) find the comovement to range between 0.07 and 0.28 and Foerster et al. (2011) document the level of average pairwise correlation to be 0.19, for instance.

The correlation process seem to have been more erratic in the earlier years. Yet, we cannot observe a clear time trend, even though the average correlation seem to have decreased slightly during 1980s. In particular, we can observe that the comovement is not constant but fluctuates considerably over time. Moreover, the conditional correlation has increased substantially during the recession periods in 1970s, 1980s and 2007-09 indicating that the sectors were comoving stronger in these downturns than otherwise. However, during the depressions in early 1990s and 2001, there is no noticeable increase in sectoral comovement in US manufacturing. This might reflect that fact that these two recessions were less related to manufacturing sector than other busts.

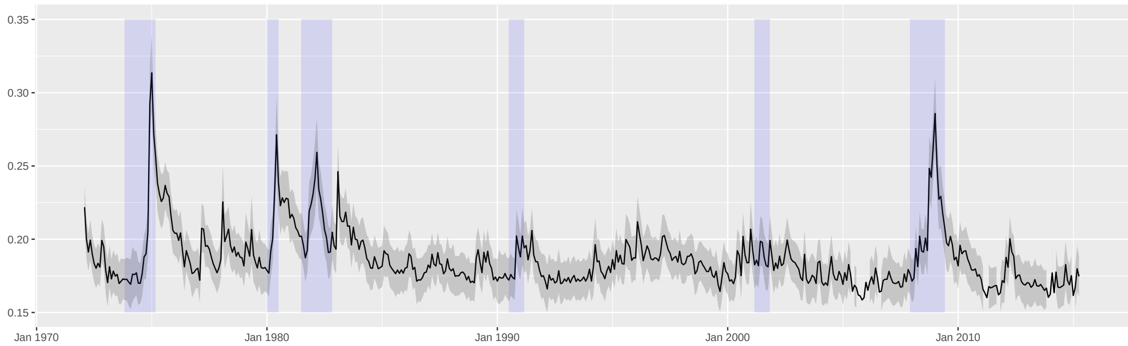

Figure 2.1: Average conditional correlation with 95 percent confidence bands and US recession periods

A further issue considering sectoral comovement that we can address with the disaggregated manufacturing data is, whether and how the dynamics differ between non-durable and durable goods sectors. Already Lucas (1977) states that the production of durable goods have more amplitude than the non-durable goods. Also Mankiw (1985) assesses the role of durable goods sectors for business cycles and points out that the sectors producing durable goods are essential for business cycle fluctuations and therefore, understanding movements in durable goods sectors is important for understanding business cycles at large. Christiano and Fitzgerald (1998) show also that the comovement of hours worked with total hours worked is in durable goods manufacturing higher than across the non-durable manufacturing sectors.

For this purpose, we calculate the average conditional correlation across durable respectively non-durable goods sectors separately. Looking at the average time-varying correlations in Figure 2.2, we can observe that the conditional correlation across sectors producing durable goods is noticeably higher and more erratic than the one of non-durable goods sectors. The slight increase in the correlations during the recession of 2007-09 is also observable among non-durable goods sectors, however, less pronounced. These findings supports the fact that durable good sectors tend to comove in general stronger than non-durable goods sectors implying stronger sectoral linkages and/or greater importance of aggregate shocks as the magnitude of common shocks is equal to all sectors. Yet, not only the level of comovement is higher but also the comovement across durable goods sectors fluctuates more suggesting that the reaction to (aggregate) shocks are alternating over time.

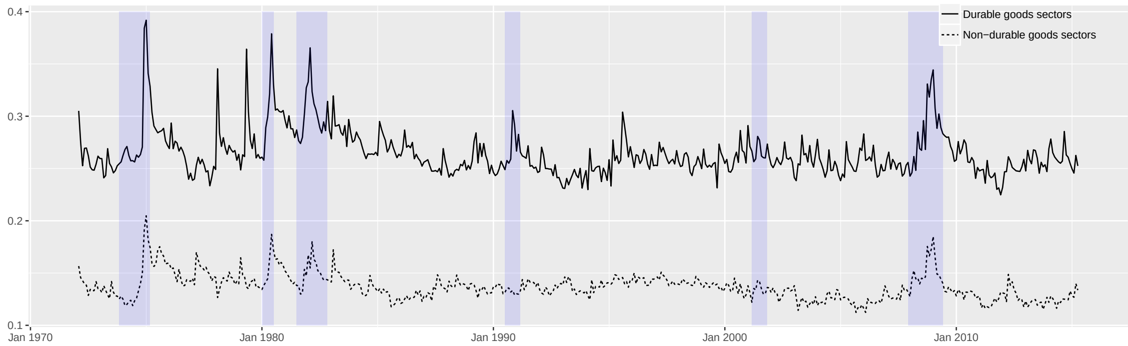

Figure 2.2: Average conditional correlations among durable goods and non-durable goods sectors and US recession periods

2.5.2 Results for Germany

We repeat the same analysis for German data. The descriptive statistic of the pairwise DCC model estimates are reported in Table 2.2. According to this table, the asymmetric version of the model is preferred only in 9 cases by the log-likelihood criterion indicating that the reactions to common factors tends to be symmetric regarding the sign of shocks for most of the sector pairs.

Table 2.2: Descriptive statistics of the bivariate DCC model estimates for German data

Statistic	N	Mean	St. Dev.	Min	Median	Max
a	209	0.040	0.062	0.000	0.013	0.362
b	209	0.744	0.317	0.000	0.912	0.998
γ	11	0.410	0.174	0.039	0.441	0.719

Figure 2.3 plots the average conditional correlation together with 95% confidence intervals and the German recession periods (defined as at least two consecutive

quarters of negative growth of real GDP). Again, in some economic downturns the correlations jump and we can observe the noticeably increase in the correlation during the economic crisis in 2007-09 also in the German data. Nevertheless, there are also some periods of higher correlations which do not overlap with German recession phases.

Altogether, the pattern of the fluctuations as well as the level of average conditional correlation are found to be quite similar both in the US and in Germany for the overlapping sample period.

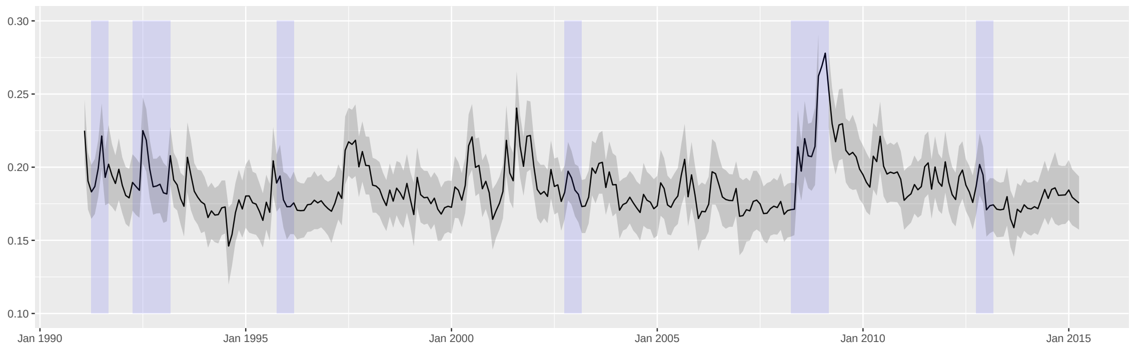

Figure 2.3: Average conditional correlation with 95 percent confidence bands and German recession periods

2.6 Explaining dynamics of sectoral comovement

In the previous chapter, we found evidence for fluctuations in the conditional correlations between monthly growth rates of IPI in manufacturing sectors both in the US and in Germany. However, with DCC-GARCH analysis we cannot assess the causes for the non-constant sectoral correlations. In general, sectors comove because of the direct effects of common factors and pass-through of shocks due to sectoral linkages. Therefore, the changes over time in interindustry comovement indicate altering magnitude of common shocks, changing sensitivity of sectors to common factors and/or varying strength of propagation of shocks. Yet, since the intensity of the pass-through depends mainly on sectoral linkages which are assumed to be rather constant or changing slowly, the first two reasons should be the main drivers of the fluctuations in interindustry correlations.

First, we examine which common factors are relevant for sectoral comovement in a regression analysis. The aggregate factors we consider include real monetary policy rates (in levels) (*IntR*) to account for monetary policy changes , stock market indices (*SP500* rsp. *DAX*) to incorporate financial market developments as well as changes in oil prices (*oil*) to control for movements in oil prices. Moreover, we also want to consider the role of volatility respectively uncertainty for the comovement since its role for real economic activity has received a lot of attention lately. The stock market volatility (*volSP500* rsp. *volDAX*) is combined series from annualized monthly standard deviation of the daily returns and implied volatility index, as implied volatility index is not available for the whole sample period. Furthermore, we include the Business Confidence Indicator (*BCI*), which is based on business tendency survey data, to account for changes in firms' confidence. Additionally, to examine if comovement increases during recessions, we add also a recession dummy respectively dummies for each recession. The data definitions and sources are reported in Tables A.7 and A.8 in A Appendix.

Table 2.3 displays the results of the estimated models for US data with the mean dynamic conditional correlation as the dependent variable. Standardized coefficient and heteroskedasticity robust standard errors are reported. In the baseline model, column (1) in Table 2.3, we include one lag of the dynamic correlation (DC_{t-1}) to account for serial correlation as well as one period lagged measures of the aggregate variables.

We find that the coefficients on stock market index and confidence indicator are significant and negative, suggesting that the higher the stock market index respectively the confidence, the lower the comovement. In other words, in times of pessimism the sectoral production decisions are more similar than in periods of high expectations. In addition, the coefficient on stock market volatility is positive and significant, even though smaller size than for *SP500* and BCI, implying higher correlations in times of high volatility. On the other hand, interest rates and oil prices seem to be unrelated for the correlations between the sectors.

Table 2.3: Estimation results of regression models for US data

	Mean dynamic correlation (DC)		
	(1)	(2)	(3)
DC_{t-1}	0.7320***	0.7320***	0.7250***
	(0.0388)	(0.0389)	(0.0411)
$volSP500_{t-1}$	0.0901***	0.0772***	0.0684***
	(0.0271)	(0.0280)	(0.0261)
$SP500_{t-1}$	−0.1120***	−0.1070***	−0.1060***
	(0.0336)	(0.0332)	(0.0299)
Oil_{t-1}	0.0208	0.0148	0.0073
	(0.0212)	(0.0219)	(0.0203)
$IntR_{t-1}$	−0.0387	−0.0382	−0.0293
	(0.0321)	(0.0318)	(0.0253)
BCI_{t-1}	−0.1110***	−0.0967***	−0.1050***
	(0.0336)	(0.0311)	(0.0383)
d.rec.US		0.0396	
		(0.0303)	
d.rec.7375			0.0328
			(0.0479)
d.rec.80			0.0348
			(0.0439)
d.rec.8182			0.0098
			(0.0312)
d.rec.9091			−0.0150
			(0.0230)
d.rec.01			−0.0011
			(0.0196)
d.rec.0709			0.0441
			(0.0395)
AIC	755.2	759.2	785.5
Observations	518	518	518
R^2	0.7700	0.7710	0.7730
Adjusted R^2	0.7670	0.7680	0.7680
Note:		*p<0.1; **p<0.05; ***p<0.01	

Next, we assess the meaning of the recession periods in model (2). The recession dummy variable (*d.rec.US*) takes 1 during US recession periods and zero otherwise. The recession dummy turn out to be positive but not significant while the other coefficients remain similar, even though slightly lower compared to the model (1). This indicates that the included common factors are able to capture to a great extent the increase in comovement during some of the recessionary periods. When examining the effects of each recession period separately (column 3), all coefficient turn out to be insignificant. Yet, while most of the dummy variables are positive, two are found to be negative indicating that during the economic slowdown in early 1990s as well as beginning of 2000s the sectoral comovement in manufacturing even might have even decreased somewhat. This suggests that these recessions were hardly related to manufacturing sector, i.e. the factors causing these specific downturns were not of great importance for manufacturing firms.

We repeat the same exercise for German data. According to the results in Table 2.4, the findings are quite different compared to results for the US. The only variable, which is found to have a significant coefficient in the model (1), is the stock market volatility. When we add the recession dummy variable (column 2), it turns out to be positive and significant implying that the included variables cannot capture fully the higher correlations during recessionary times. The results of the model with separate dummy variables for each recession (column (3)) reveal that correlation between manufacturing sectors did increase during some recession periods but not always as was found also for the US data. However, given that the sample period is much shorter for Germany than for the US, the comparison of the results is not straightforward.

In the next step, we study if the movements in the correlations could be due to changes in the variance of common factors. For this purpose, we calculate the rolling standard deviations (12 months) of the US variables used in the regression analysis since the sample period for the US is much longer. The results in Figure 2.4 implicate that the variances of the common factors do fluctuate considerably. While the variability of interest rates has reduced over time, the plots for the other variables look quite different. The variances peak for all factors expect interest rates during the financial crises around 2007-2009. Given that the results of rolling analysis are greatly influenced by the window length, we do not want to emphasize these findings too much but exploit them to point out that the variances of the aggregate factors are clearly non-constant over time.

Table 2.4: Estimation results of regression models for German data

	Mean dynamic correlation (DC)		
	(1)	(2)	(3)
DC_{t-1}	0.6680***	0.6500***	0.6390***
	(0.0525)	(0.0515)	(0.0491)
$volDAX_{t-1}$	0.2960***	0.2910***	0.2130*
	(0.0974)	(0.1020)	(0.1170)
DAX_{t-1}	0.0355	0.0789	0.0899
	(0.0618)	(0.0611)	(0.0601)
Oil_{t-1}	0.0948	0.0423	−0.0237
	(0.0617)	(0.0578)	(0.0551)
$IntR_{t-1}$	0.0335	−0.0077	−0.0705
	(0.0631)	(0.0626)	(0.0640)
BCI_{t-1}	−0.0517	−0.0483	−0.0704
	(0.0533)	(0.0516)	(0.0522)
d.rec.GE		0.1180**	
		(0.0565)	
d.rec.91			0.0884*
			(0.0482)
d.rec.9293			0.0981*
			(0.0526)
d.rec.9596			−0.0217
			(0.0215)
d.rec.0203			0.0151
			(0.0399)
d.rec.0809			0.1540*
			(0.0802)
d.rec.1213			−0.0158
			(0.0661)
AIC	469.5	469.1	488.3
Observations	225	225	225
R^2	0.6090	0.6200	0.6350
Adjusted R^2	0.5980	0.6070	0.6140
Note:	*p<0.1; **p<0.05; ***p<0.01		

Finally, to address the possible changes in the sensibility of the sectors to common factors, rolling regression analysis of the model (1) for the US is conducted. We set the window length to 8 years. Looking at the rolling standardized coefficients in Figure 2.5, we observe that they exhibit considerable movements. The coefficient of the lagged dependent variable (DC_{t-1}) decrease continuously until 2008 and increase then sharply back to the higher level. The coefficient on stock market volatility is found to have increased since mid 2000s as well as the coefficient on business confidence index has become also more negative since early 2000s implying greater influence of these variables on the sectoral comovement during the last decades. Yet, the estimates for stock market index, oil prices and interest rates do even change sign during the sample period indicating unstable relationships to intersectoral correlations.

In summary, our findings suggest that both the time-varying variance of the common factors as well as the changes in the sensitivity of sectors to aggregate factors play role for the movements in sectoral comovement. Yet, further research is needed to explore these channels in more detail and to account, for instance, for the possible time-varying interdependencies between the common factors itself.

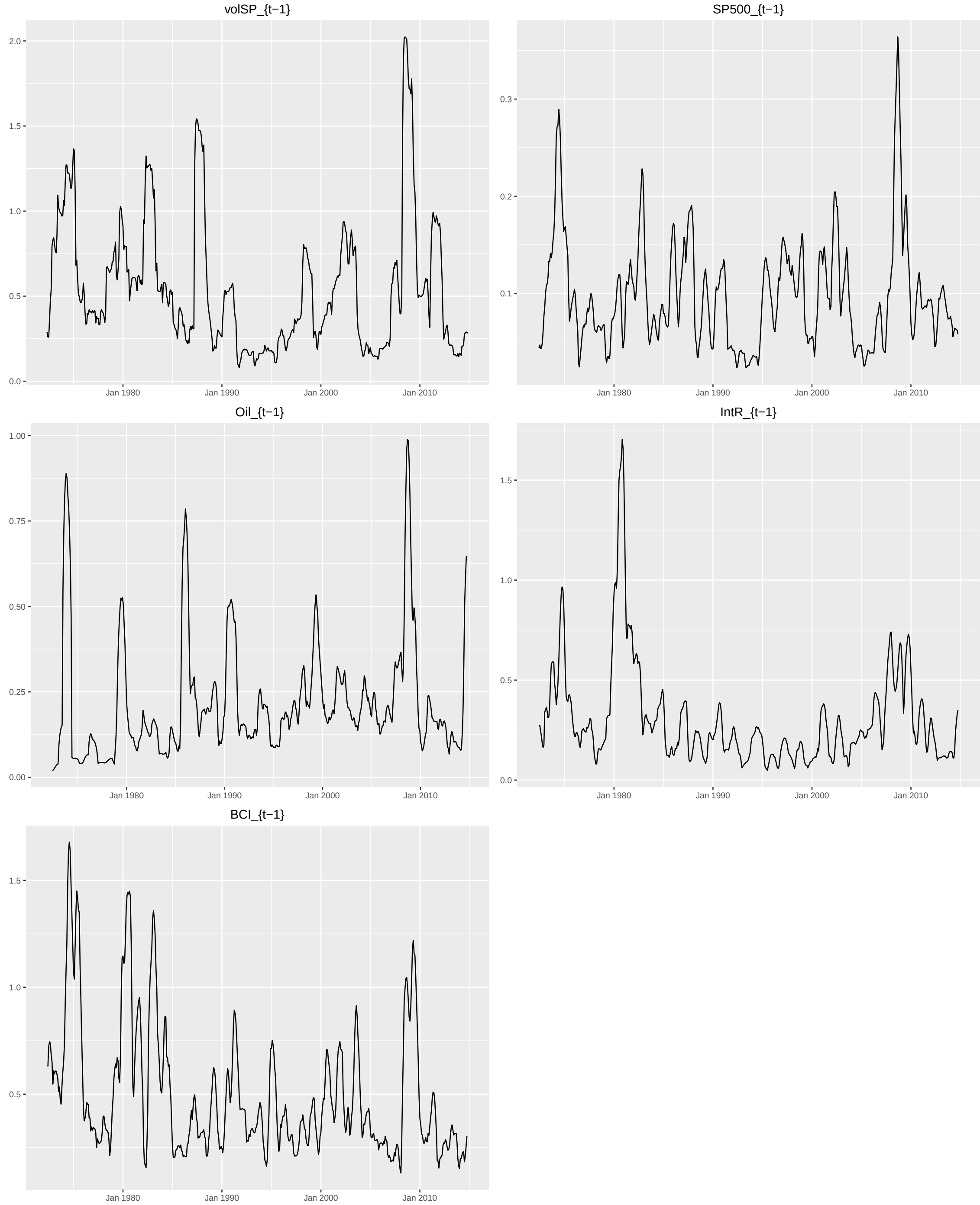

Figure 2.4: Rolling standard deviations (window length of 12 month)

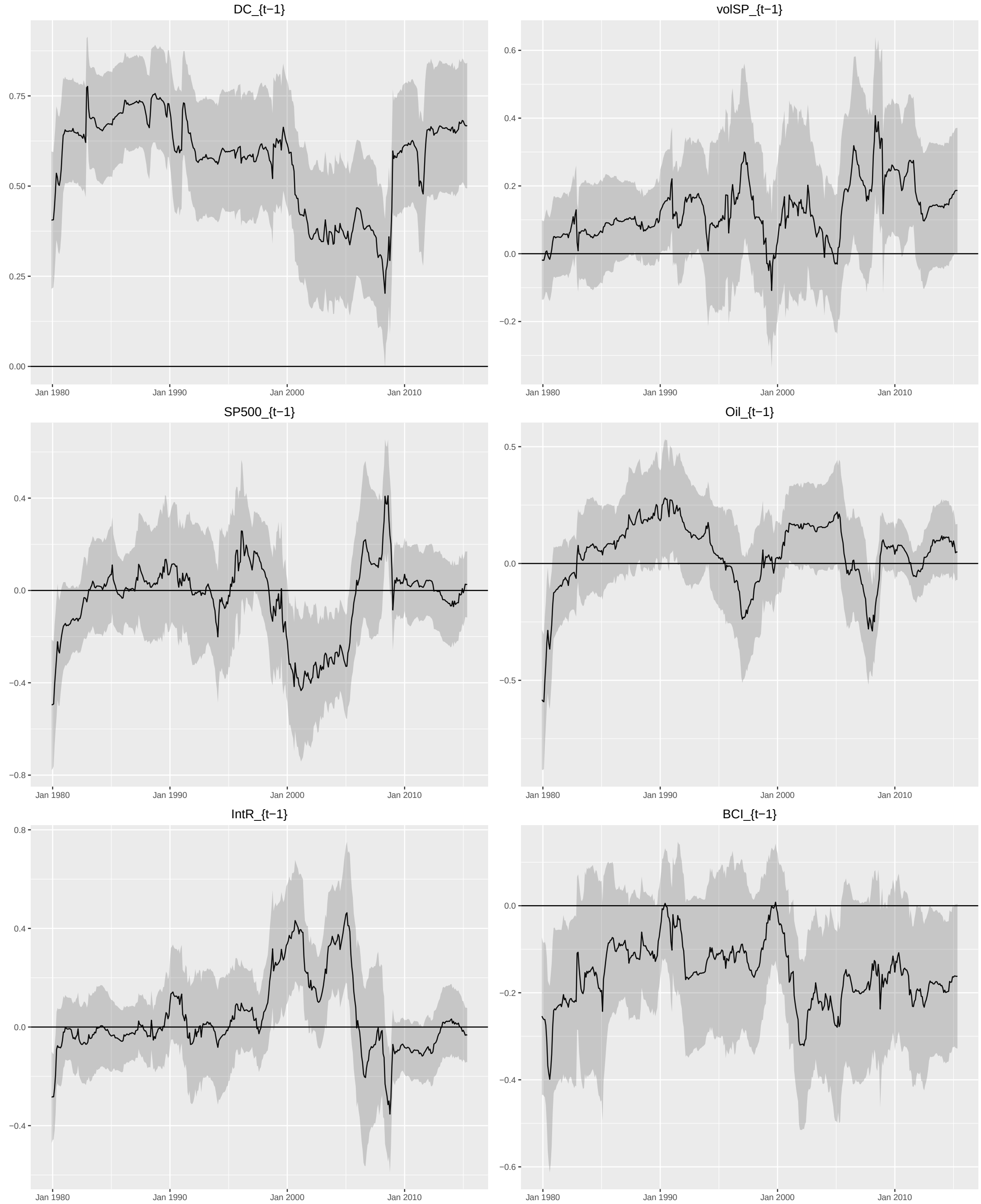

Figure 2.5: Rolling standardized regression estimates

2.7 Conclusions

In this paper, we examine dynamics of sectoral comovement among disaggregated manufacturing sectors in the US and in Germany. Given that the aggregate volatility depends not only on the volatility of individual sectors but to a great extent of correlations between them, understanding sectoral comovement is crucial for explaining business cycles. For this purpose, we employ a multivariate dynamic conditional correlation (DCC) general autoregressive conditional heteroskedasticity (GARCH) framework to assess the time-varying and asymmetric aspects in volatilities as well as in correlations of growth rates of manufacturing production. The advantage of DCC-GARCH is that we can examine possible changes in conditional correlations depending on state or circumstances of the economy as well as generally over time.

Our results reveal that variance of industrial production growth in manufacturing sectors in the US tend to be asymmetric i.e. most industries react stronger to unfavourable than favourable shocks. However, we find only limited evidence for asymmetries in correlations what implies that most sector pairs react similarly to common shocks, irrespective of the sign of shocks. Most notably, we show that the sectoral comovement in the US and German manufacturing is not constant but shows irregular movements. Interestingly, the correlations seem to have increased considerably during some economic downturns, especially during the recession in 2007-09, but there are also recessionary periods in which the comovement hardly changes. This indicates that contractions tend to be more synchronized than expansions in manufacturing sector.

Since sectors comove because of the direct effects of common factors and pass-through of (aggregate and sectoral) shocks, the changes over time in interindustry comovement indicate altering magnitude of common shocks, changing sensitivity of sectors to common factors and/or varying strength of propagation of shocks. Yet, since the intensity of the pass-through depends mainly on sectoral linkages which are assumed to be rather constant or changing slowly, the first two reasons should be the main drivers of the fluctuations in interindustry correlations.

We examine first which aggregate factors are relevant for sectoral comovement in a regression analysis. Our findings for the US indicate that stock market index as well as business confidence are related to sectoral comovement. Especially, in times of pessimism the sectoral production decisions are more similar than in periods of high expectations. Also stock market volatility is found to play a role implying higher correlations in times of high volatility. Yet, for Germany, only the stock market volatility is found to have a significant impact on sectoral correlations. Furthermore, our findings suggest that time-varying variance of common factors as well as the altering impact of these factors is relevant for the fluctuations in sectoral comovement. However, further research is needed to explore these issues in more detail.

Chapter 3

The role of common expectations for sectoral comovement: evidence from business survey data

3.1 Introduction

The term business cycle refers to two types of comovement: aggregate and sectoral comovement. The former describes the synchronized movements of aggregate variables such as consumption, investment and hours worked. The latter outlines the comovement of inputs and outputs across sectors. The joint behavior of sectors is commonly interpreted as evidence for aggregate shocks causing the business cycle (e.g., Lucas (1977)), such as monetary policy or technology shocks affecting all industries in a similar manner. However, generating sectoral comovement in business cycle models has turned out to be difficult even in case of aggregate technology shocks as discussed by Christiano and Fitzgerald (1998), for instance. Moreover, the finding of Hornstein (2000) that production is stronger correlated across sectors than productivity, also referred to as *excess comovement*, do not support an aggregate technology shock based explanations. Further common factors that could account for sectoral correlations are hard to identify.

Yet synchronized movements across sectors do not necessarily need to be a result of common shocks but also sector-specific shocks could be propagated through sectoral linkages to other sectors causing industries to move together. Nevertheless, sectoral linkages tend not to be strong enough, especially at the more detailed level of disaggregation, to account for a realistic level of interindustry comovement (e.g., Foerster et al. (2011)). As production complementarities have not turned out to be a successful explanation, further ideas have been proposed. Christiano and Fitzgerald (1998) discussed *information externalities* as a potential source of interindustry comovement and Veldkamp and Wolfers (2007) were the first to offer a theoretical model for the role of common expectations, or aggregate information, for sectoral comovement. They argue that complementarity in information acquisition could also explain the excess comovement puzzle - an important but often ignored empirical finding. Their explanation is based on the idea that firms collect also aggregate information about productivity developments, as it is cheaper to acquire than sector-specific information. Based on similar anticipations, firms tend to make similar input and output decisions even though their actual productivity developments may differ.

As noted by Veldkamp and Wolfers (2007) themselves, the challenge with their hypothesis is that it is difficult to assess empirically since the anticipations, on which agents are assumed to base their decisions, are not directly observable. Data on expectations, and especially data on expectations about productivity developments, is scarce. Not to mention that sector-level measures of expectations are hardly available. Lamla et al. (2007) study the meaning of information complementarities for sectoral comovement by examining if newspaper news affect firms' perceptions and expectations. Their results support the hypothesis Veldkamp and Wolfers (2007) as they find that economy-wide news affect firms' business assessments and plans more than sector-specific media information does.

In this paper, we take a different approach and try to link the common expectations directly to the level output comovement across manufacturing sectors. For this purpose, we extract the common component of sectoral expectations and examine how much of the correlation between disaggregated manufacturing output it can explain. Moreover, to reveal how unexpected changes in common expectations are propagated through the economy, it is important to know the magnitude of the responses in different sectors. To tackle these issues, we construct in this paper novel measures of anticipated labour productivity changes at both sector as well as aggregate levels to assess this hypothesis empirically. The proposed measures of expected labour productivity changes are based on micro data of business tendency survey.

In Section II, we discuss the theoretical background and related literature in more detail. Section III describes the data from business tendency surveys (BTS) we use in this paper. Our data source is the monthly Swiss manufacturing survey of KOF Swiss Economic Institute. Firms in the manufacturing sector are surveyed each month and asked, among other, questions about their expectations. Using the question Q.8b "Expectations about production for the next three months" and Q.8d "Expectations about future number of employees for the next three months" we are able to define unique, direct measures of anticipated (labour) productivity changes. Furthermore, we show that our measure of anticipated productivity changes effectively tracks labour productivity.

With our sector level expectations measures, we explore the hypothesis of Veldkamp and Wolfers (2007) empirically in Section IV and shed light on the role of aggregate information for inter-industry synchronization. To find out whether there is such an aggregate information component in sectoral expectations and to discover the industry-specific sensitivity, we employ factor analysis. The results of this exercise support the hypothesis of Veldkamp and Wolfers (2007) as we find that the firms in different sectors actually do base their expectations on similar information. To assess the quantitative meaning of common information for sectoral comovement, we examine how much of inter-industry correlations of manufacturing production are related to the common factor of the sectoral expectations. We show that after accounting for the aggregate information, the pairwise correlations across manufacturing industries are considerably lower. This implies that information complementarities are relevant for explaining the synchronized movements across industries.

Moreover, through VAR analysis we can observe how the dynamic effects of manufacturing production to a surprise change in the common component of expectations differ across sectors. Our findings reveal that there is a great deal of heterogeneity in the reactions of the disaggregated manufacturing sectors to changes in expectations. In particular, expectation shocks are found to propagate mainly through the sectors producing capital goods and consumer durables. This is line with the expenditure side literature on news shocks (e.g., Beaudry et al. (2011) and Beaudry and Portier (2014)) as it is theoretically argued and empirically found that investment reacts most to information about future fundamentals. Furthermore, the responses of durable and non-durable consumption goods differ considerably as consumer non-durables seem to be almost unrelated to firms' expectations.

Section V contains conclusions.

3.2 Theoretical background and related literature

The term *business cycle* refers, on the one hand, to joint movement of aggregate variables such as output, investment and consumption and on the other hand, to synchronized movements of disaggregated sectoral variables. Therefore not only *aggregate comovement*, but also *sectoral comovement*, is a main feature of business cycles. However, business cycle research has so far chiefly concentrated on examining comovement between aggregate variables. This is likely the case as standard business cycle models assume only one economic sector. Obviously, within such a framework, sectoral comovement cannot be addressed. Additionally, within two-sector models (with investment and consumption good sectors) generating sectoral comovement has turned out to be difficult even in the case of aggregate technology shocks as discussed by Christiano and Fitzgerald (1998), for instance. This is because investment increases more than consumption in response to a technology shock and, therefore, labour should move from the consumption sector to the investment sector. This implies negative comovement of hours worked in these two sectors.

Yet the synchronized movement across sectors does not necessarily need to be a result of common shocks but also sector-specific shocks could be propagated through the other sectors causing industries to move together. The necessary condition for this kind of propagation is a mechanism which limits the law of large numbers, that is, that sector-specific (uncorrelated) disturbances do not cancel out. Long and Plosser (1983) show that in a multisector framework with intersectoral linkages (input-output dependencies), independent sectoral shocks alone may induce sectoral comovement. However, they assume full depreciation of capital and when this assumption is relaxed, many properties no longer generalize as noted by Rebelo (2005). Similarly Horvath (1998, 2000), Dupor (1999), Shea (2002) and Carvalho (2008) examine the role of intersectoral dependencies transmitting sector-specific shocks within structural multisector models. Nonetheless, sectoral linkages tend to be too weak, especially at more detailed level of disaggregation, in order to account for a realistic level of interindustry comovement (see, Foerster et al. (2011)).

Further, these models usually assume contemporaneous sectoral linkages as they are based on data from yearly input-output tables. Yet, the more realistic assumption is that the intermediate goods as well as capital goods are produced at least a period (month or quarter) earlier before they flow into production of other goods, that is, production complementarities tend to operate with lag(s). In such a framework, it is much harder to make sectors to move synchronized. Many of the empirical applications employ yearly data and legitimate the assumption of contemporaneous linkages. In addition, further mechanisms such as inventories (Cooper and Haltiwanger (1990)) and trade credit (Raddatz (2010)) have been proposed, playing though at best only a moderate role in explaining synchronized sectoral movements.

A further proposition how sectoral comovement might evolve is based on information complementarities or common information. Christiano and Fitzgerald (1998) discuss *information externalities* as a potential source of inter-industry comovement. However, Veldkamp and Wolfers (2007) were the first to offer a theoretical model for the role of expectations, or aggregate information, for sectoral comovement. However, the idea that changes in expectations might be an essential driving force of fluctuations in aggregate economic activity was brought to light already at the beginning of the 20th century (see, Beveridge (1909), Clark (1934) and Pigou (1927)). For long time, this idea did not receive much attention as the tradional explanations of business cycles concentrated on productivity shocks. Beaudry and Portier (2004, 2006) combine the ideas and argued that changes in expectations reflect news about future fundamentals, i.e., technological changes are anticipated such that they become known in the form of news before they are actually realized in a higher technological level. As already suggested in the early 90s by Fama (1990), Beaudry and Portier (2006) use stock prices to identify the changes in expectations of economic agents and markets. Since then, this hypothesis has received growing attention (e.g., Jaimovich and Rebelo (2009), Beaudry and Lucke (2010), Schmitt-Grohe and Uribe (2012), Forni et al. (2014)). The main question in this literature is, whether these kind of news are important for business cycles through their impact on firms' motivation to start to invest immediately in anticipation of future demand and build up capital goods before any changes in fundamentals have actually occurred. The higher demand for investment should in turn lead to an increase in consumption generating an aggregate expansion. In general, these models are attractive as they can generate recessions without technological deterioration but as a reaction to over-optimism.

The news literature, like the business cycle literature in general, has mainly focused on explaining and documenting aggregate comovement between macroeconomic aggregates such as investment and consumption. However, if news shocks are important for business cycles, they should also be relevant for sectoral comovement, i.e., be able to explain the synchronized movements across disaggregated sectors. Christiano and Fitzgerald (1998) argue that anticipations of future developments of the economy are incorporated in firms' current investment decisions. Due to information externalities, a firm is interested in actions of other firms because these actions may reveal information on a topic, such as the future state of the economy, which is also of interest for the firm itself. Further, Veldkamp and Wolfers (2007) note that complementarity in information acquisition could also explain the excess comovement puzzle - documented by Hornstein (2000) i.e., why production is stronger correlated across sectors than productivity. This stylized fact does not support the idea of an aggregate productivity shock as a driver of economic fluctuations. Veldkamp and Wolfers (2007) argue that firms try to economize their information costs and therefore, gather also aggregate information, as it is cheaper to acquire than industry-specific information is. From this it follows that sectoral expectations tend to be positively correlated and thus, also the production choices based on these expectations.

More precisely, Veldkamp and Wolfers (2007) propose an island model of production in which the productivity of island i is unknown at the start of each period and therefore, information or expectation about productivity matter. Each sector's productivity is given by

$$z_i = \mu_z + \beta_i \bar{z} + \eta_i + e_i \tag{3.1}$$

where μ_z is a known aggregate component, $\bar{z}$ is an unknown but learnable aggregate component, η_i is an unknown but learnable industry-specific component, e_i is an unknown and unlearnable industry-specific component, and $\bar{z} \sim N(0, \sigma_z^2)$, $\eta_i \sim N(0, \sigma_n^2)$ as well as $e_i \sim N(0, \phi_j^2)$. β_i measures the strength of which sector i's productivity comoves with the aggregate.

The industry-specific signal about the industry i's (future) productivity is defined by:

$$s_i = \beta_i(\bar{z} + e_0) + \eta_i \tag{3.2}$$

where $\bar{z} + e_0$ is the signal about aggregate productivity, η_i is the sector-specific signal and $e_0 \sim N(0, \phi_0^2)$. Thus, the industry-specific signal is the sum of the (noisy) aggregate signal and industry-specific signal. The expectations about z_i, $E[z_i|I_i]$, are the driving force of this model as the labour effort is a function of anticipations about productivity and production as well, since it depends on labour input:

$$y_i = z_i n_i \tag{3.3}$$

$$n_i^* = \frac{1}{\rho Var[z_i|I_i]}(E[z_i|I_i] - \psi) \tag{3.4}$$

where n^* is the first-order condition for labour, I_i denotes the available information.

The empirical evidence for the role of anticipations for synchronized movements across industries in general and for excess comovement in particular is still scarce. Lamla et al. (2007) study the meaning of information complementarities for sectoral comovement as suggested by Veldkamp and Wolfers (2007) by examining if newspaper news affect firms' perceptions and expectations. Their results support the hypothesis Veldkamp and Wolfers (2007) as they find that economy-wide news affect firms' business assessments and plans more than sector-specific media information does.

3.3 Data Description

Our main data source is the monthly KOF manufacturing survey for Switzerland. That is a voluntary monthly survey among firms in the manufacturing sector. In this analysis, we use the answers of the following two questions: Q.8b "Expectations about production for the next three months" and Q.8d "Expectations about future number of employees for the next three months". The answer categories are "increase", "stay at same" and "decrease". The anticipated future changes are defined through a contingency table (Table 3.1) where $P(.,.)$ denotes the relative frequency of certain expectation combination. The expectation at period t about future production at period $t+1$ are denoted by $E_t y_{t+1}$ and about number of employees by $E_t n_{t+1}$.

Table 3.1: Expectation combinations

	$E_t y_{t+1}$ = Increase	$E_t y_{t+1}$ = Unchanged	$E_t y_{t+1}$ = Decrease
$E_t n_{t+1}$ = Increase	P(+,+)	P(+,=)	P(+,-)
$E_t n_{t+1}$ = Unchanged	P(=,+)	P(=,=)	P(=,-)
$E_t n_{t+1}$ = Decrease	P(-,+)	P(-,=)	P(-,-)

We interpret that a firm that expects an increase in production but unchanged or decreasing number of employees, anticipates a positive (labour) productivity increase. The same argumentation goes the other way round for negative changes. Obviously, it is not guaranteed that the both variables are assumed to change with same intensity such that also in the case both are expected to increase, output might increase more than the number of employees. Yet, we do not examine this issue here further, but concentrate on the cases which different directions are assumed. Moreover, we weight the answers by the capped number of employees[1].

Following Kawasaki and Zimmermann (1986) we calculate from this frequency distribution a measure of anticipated productivity changes as a weighted balance of positive and negative anticipated changes. We try two different specifications:
$EXP.B1 = (APPC - ANPC)/(APPC + ANPC)$ where
Anticipated positive productivity changes $(APPC) = P(=,+) + P(-,+) + P(-,=)$
Anticipated negative productivity changes $(ANPC) = P(+,=) + P(+,-) + P(=,-)$
as well as $EXP.B2 = (APPC - ANPC)/(APPC + ANPC)$ where
Anticipated positive productivity changes $(APPC) = P(=,+) + P(-,+)$
Anticipated negative productivity changes $(ANPC) = P(+,=) + P(+,-)$.

[1] To hinder that big firms dominate the results, a size cap is often applied in business tendency surveys. In our case, the threshold is set to 500, that is, for $w > 500$ where w is the number of full time equivalents employees, the weight is then calculated by $w^{0.7}$ instead just using w.

All measures are seasonally adjusted and normalized. As the question about the employment expectations was introduced only in 2004 for the monthly survey, our sample period is January 2004 - June 2017. As many economic variables are available only at a quarterly frequency, we aggregate our monthly measures also to quarterly variables taking quarterly averages. Further, we also construct sector level measures for the NACE letter code level industry branches[2].

Moreover, we use the data on manufacturing production which we take from the Swiss Federal Statistic Office (SFSO). We adjust the series seasonally in order to be able to use the quarter-on-quarter growth rates. Also the value added and employment data come from the SFSO. We calculate labour productivity as value added divided by labour force. The relative sizes of the dissaggregated sectors we derived from sectoral turnovers also provided by SFSO.

3.3.1 Firms' expectations and labour productivity

Since we argue that our measures could capture changes in labour productivity, we first visually explore the relationship between these variables. We do not expect that our expectations measures are able to track all the movements in labour productivity, especially the transitory movements which are hard to be anticipated, but they should be able to track the overall trend. Figure 3.1 plots quarterly averages of EXP.B1, EXP.B2 and the HP-trend of labour productivity growth (to illustrate trend growth). The positive comovement between EXP.B2 and labour productivity growth is obvious. We can observe that the expectations started deteriorate slowly during 2007 before dropping considerably end of 2008 and 2009. However, the anticipations recovered also quickly showing a sharp increase.

[2]Definition of letter classes: Food (CA = NACE 10-12), Textiles (CB = 13-15), Wood (CC = 16-18), ChemicalsPharma (CD + CE + CF = 19-21), Rubber (CG = 22-23), Metals (CH = 24-25), Electronics (CI = 26), Electricals (CJ = 27), Machinery (CK = 28), Transport (CL = 29-30), Other (CM = 31-33). The difference to the official letter classes of NACE is the class CN which combines the CD+CE and CF classes. As the number of observations is very small for the class CL in the survey data, we cannot calculate the sector level expectations measure for this class.

Yet, the EXP.B1 measure does not seem to be able to capture the changes in productivity so well. Thus, we will use the $EXP.B2$ measure in this analysis as it is found effectively to reflect information about changes in labour productivity.

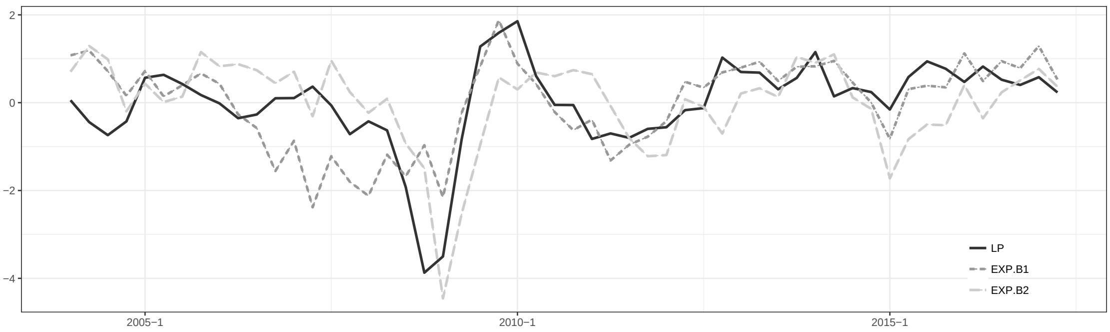

Figure 3.1: EXP.B1, EXP.B2 and labour productivity growth (scaled)

3.4 Aggregate versus sectoral information

In this section, we analyse the hypothesis of Veldkamp and Wolfers (2007) with our sector level indicators in more detail and shed light on the role of aggregate information for sectoral production movements. First, we document sectoral comovement as well as excess comovement in Swiss data. The correlation matrix between manufacturing production growth in disaggregated sectors[3] in Table 3.2 illustrates the amount of comovement. The correlations vary between -0.32 and 0.6 indicating a rather heterogeneous structure of intersectoral relations. The mean correlation amounts 0.21 implying that industries move on average moderately together which is in line with the findings of Long and Plosser (1987) and Foerster et al. (2011).

[3]Due to data availability, we use following letter-level sector groups: Food (CA = NACE 10-12), Textiles (CB = 13-15), Wood (CC = 16-18), Chemicals (CD+CE = 19-20), Pharma (CF = 21), Rubber (CG = 22-23), Metals(CH = 24-25), Electronics (CI = 26), Electricals (CJ = 27), Machinery (CK = 28), Transport (CL = 29-30), Other (CM = 31-33)

Table 3.2: Correlations between sectoral manufacturing production growth

	1	2	3	4	5	6	7	8	9	10	11
Food											
Textiles	0.35										
Wood	0.26	0.35									
Chemicals	-0.32	-0.03	0.07								
Pharma	0.15	0.23	0.09	-0.01							
Rubber	0.05	0.12	0.39	0.23	-0.08						
Metals	0.14	0.20	0.55	0.15	0.01	0.58					
Electronics	0.13	0.36	0.46	-0.09	0.23	0.42	0.31				
Electricals	0.24	0.21	0.39	0.11	0.26	0.27	0.24	0.27			
Machinery	0.26	0.26	0.39	-0.02	0.09	0.17	0.42	0.20	0.35		
Transport	-0.01	0.24	0	-0.12	0.07	0.36	-0.01	0.13	0.29	0.16	
Other	0.18	0.30	0.60	0	0.10	0.27	0.37	0.50	0.17	0.50	0.11

To illustrate excess comovement, we calculate the average pairwise correlations for growth rates of industrial production as well as labour productivity for the manufacturing sectors. In Figure 3.2 we can observe that output is in general higher correlated across sectors than productivity is as originally found by Hornstein (2000). Again, this indicates that traditional productivity shocks do not appear to be the main reason for the synchronized movements.

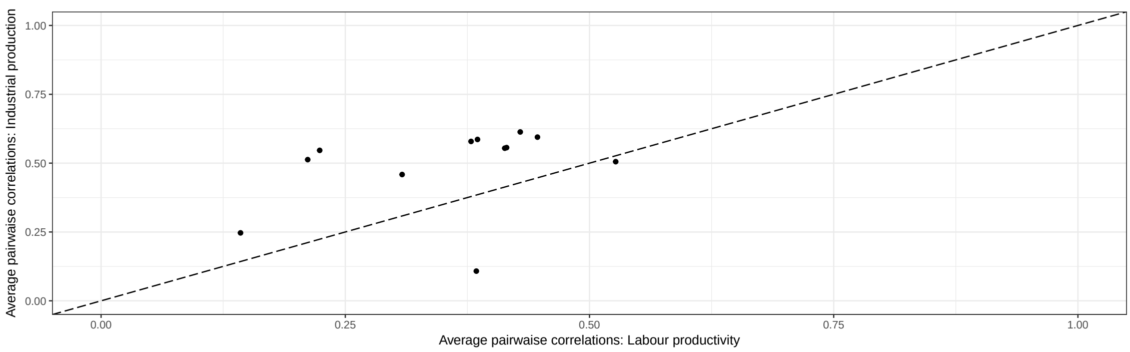

Figure 3.2: Excess comovement in Swiss Data

The descriptive statistics of the manufacturing production series (in log differences) are reported in Table 3.3. The sectoral volatilities are found to differ greatly. Panel B underlines the stylized fact that the durable consumer goods sector exhibits higher volatility than non-durable consumer goods. Also capital goods production seems to be more variable than non-durable consumer goods production what is in line with the fact that investment is more erratic than consumption, especially non-durables consumption. These differences in volatilities imply that sectors are affected by different sectoral shocks or respond differently to aggregate shocks. In particular, the industries with higher volatility are presumably subject to more sector-specific disturbances.

Table 3.3: Descriptive Statistics of sectoral manufacturing production series

Panel A. Letter-code classification						
	Mean	Std. Dev.	Min.	Max.	Range	Corr.
Food	0.20	1.42	-2.63	3.23	5.85	0.13
Textiles	-0.22	4.53	-17.68	16.08	33.77	0.24
Wood	-0.18	1.01	-2.92	1.56	4.48	0.32
Chemicals	-0.10	5.22	-25.23	9.07	34.30	-0.00
Pharma	1.72	3.01	-4.76	9.32	14.08	0.10
Rubber	-0.07	2.40	-6.91	5.19	12.09	0.25
Metals	-0.06	1.82	-4.71	3.02	7.73	0.27
Electronics	0.59	3.62	-8.62	7.40	16.02	0.26
Electricals	1.43	2.96	-3.93	12.85	16.78	0.25
Machinery	-0.05	4.70	-9.47	10.01	19.48	0.25
Transport	0.60	7.69	-18.44	20.18	38.62	0.11
Other	0.46	2.06	-5.11	3.60	8.71	0.28
Panel B. MIG classification						
	Mean	Std. Dev.	Min.	Max.	Range	Corr.
Int	0.19	2.05	-12.19	2.97	15.16	0.34
Cap	0.37	2.60	-8.81	5.49	14.30	0.32
Dur	0.45	4.10	-10.68	7.68	18.35	0.24
NDur	0.81	1.77	-3.06	4.50	7.56	0.22

Notes: Corr. = Average pairwise correlations. Int = Intermediate goods, Cap = Capital goods, Dur = Durable consumer goods, NDur = Non-Durable consumer goods

As explained earlier, Veldkamp and Wolfers (2007) assume that firms pursue to collect aggregate information about productivity developments since it is cheaper to acquire than sector-specific information. Since firms need to make most input and output decisions *before* technological level (or demand) has realized and as their expectations are correlated through the public information, their input and output decisions are more synchronized than their actual productivity developments.

To find out if there is such an aggregate information component in sectoral anticipations, we employ factor analysis. We extract the common component $\bar{z} + e_0$ of sectoral expectations measures by dynamic factor analysis and to estimate the sector-specific sensitivity (β_i) to the common information component. Thus, we consider a simple model with one common trend only:

$$\mathbf{y_t} = \mathbf{Z x_t} + \mathbf{v_t} \text{ where } \mathbf{v_t} \sim \mathbf{MVN(0, R)} \tag{3.5}$$

$$\mathbf{x_t} = \mathbf{x_{t-1}} + \mathbf{w_t} \text{ where } \mathbf{w_t} \sim \mathbf{MVN(0, Q)} \tag{3.6}$$

where $\mathbf{y_t}$ is a vector of observations of sectoral expectations, $\mathbf{x_t}$ is the common trend and $\mathbf{Z}$ is a vector of factor loadings. We assume the covariance matrix $\mathbf{R}$ to be diagonal but unequal.

Table 3.4: Factor loadings of sectoral expectations

	Food	Textiles	Wood	ChemicalsPharma	Rubber	Metals	Electronics	Electricals	Machinery	Other
coef	0.36	0.25	0.38	0.29	0.51	0.46	0.46	0.17	0.51	0.36
se	0.08	0.09	0.09	0.09	0.08	0.08	0.09	0.09	0.09	0.09

The estimated factor loadings (β_i's) with their standard errors are reported in Table 3.4. Since the β_i's are found to be positive and relatively high, this indicates that sectors actually do base their expectations on similar information, yet to varying degree. The results reveal that expectations of Rubber, Electronics and Machinery industries co-vary most strongly with the aggregate signal. Electricals and Textiles sectors exhibit the lowest relation to the learnable aggregate component, which indicates that this sector comoves only weakly with the aggregate productivity development. Furthermore, the share of the variance explained by the common factor is documented in Table 3.5. Especially, the expectations of Rubber, Electronics, Machinery and Metals are to a great extent explained by the common component. All together, these results are in line with the findings of Lamla et al. (2007) and support the hypothesis of Veldkamp and Wolfers (2007) that firms seem to gather notably aggregate information.

Table 3.5: Share of variance of sectoral expectations explained by the common factor

Food	Textiles	Wood	ChemicalsPharma	Rubber	Metals	Electronics	Electricals	Machinery	Other
0.30	0.14	0.32	0.18	0.58	0.47	0.48	0.07	0.58	0.30

Next, we ask the question how much of the comovement of manufacturing production across the disaggregated sectors can common information explain? To examine this, we regress the sectoral manufacturing production series against two autoregressive lags and current and lagged values of the common component (extracted from sectoral expectations in the previous exercise). The residuals of these regressions will show the movements which cannot be explained by the common information. Therefore, the level of correlation between these residuals (compared to the correlation between actual sectoral series) indicates how much of the comovement between sectors can be accounted to common expectations. That is, the lower the correlation between these residuals, the greater the role of aggregate information for synchronized movements. Table 3.6 presents the correlation matrix. In general, the sectoral manufacturing correlations are considerably lower after accounting for the common expectation component. Nevertheless, there are great differences across the sector pairs. For some industry pairs the correlations seem to be even higher after controlling for aggregate information. All together, the average pairwise correlation is found to amount to 0.06 (compared to 0.15 between the AR-innovations) indicating that almost two thirds of output comovement can be attributed to the common component of sectoral expectations. This finding emphasizes that aggregate information plays an essential role in synchronized industry fluctuations. On the other hand, also some other aggregate factors or production complementarities seem to be meaningful since not all of the comovement can be explained through common information.

Moreover, we examine how unexpected changes in the common information component are actually propagated through the economy. To analyse how these dynamic effects differ across sectors, we estimate bivariate VARs (with 2 lags and a constant) consisting of the common expectation component and a sectoral manufacturing production variable (in log differences). Here, we use the MIG classification of the manufacturing sectors (instead of the letter-level classification) to get a clearer picture.

Table 3.6: Correlations between sectoral manufacturing production innovations after accounting for common information

	1	2	3	4	5	6	7	8	9	10	11
Food											
Textiles	0.21										
Wood	0.21	-0.04									
Chemicals	-0.18	-0.04	0.05								
Pharma	0.16	0.37	0.07	0.03							
Rubber	0.13	-0.15	0.14	0.27	0.03						
Metals	-0.03	-0.06	0.20	0.11	-0.14	0.40					
Electronics	0.07	0.17	0.10	-0.11	0.27	0.31	0.01				
Electricals	0.13	-0.08	-0.03	0.25	0.24	0.08	-0.02	-0.05			
Machinery	0.05	-0.06	0.06	-0.10	0.17	0	0.01	0.01	0.12		
Transport	0.07	0.22	-0.04	-0.11	0.13	0.13	-0.19	0.10	0.21	0.17	
Other	0.22	0	0.37	0.03	-0.02	0.10	0.06	0.26	-0.21	0.23	-0.04

Figure 3.3 displays the cumulative impulse responses of manufacturing production at the sectoral level to innovations in the common factor of expectations. The capital goods producing sector is found to react strongest showing a slow-building and apparently permanent response. A one-standard-deviation innovation to the common factor leads to roughly 4 percent higher production in capital goods sectors which indicates a sizable effect of expectation shocks. This is in line with our finding that expectations of industries like Machinery and Metals load strongly on the common information component. Furthermore, as capital goods are needed for investments and investment is theoretically argued and empirically found to react the most to information about future fundamentals, see e.g., Beaudry et al. (2011) and Beaudry and Portier (2014), our findings fit in the same picture with the results of news literature on aggregate comovement. Also the manufacturing production of durable goods reacts to a surprise change in the aggregate factor similar to the capital goods production, though somewhat weaker and slower. Nevertheless, the non-durable goods industry shows hardly any response to innovations in the common factor of expectations. This is plausible as durable consumption is known to be more cyclical and responsive to sentiments than expenditure on non-durable goods

(see, Fuhrer (1993)). The intermediate goods sector shows qualitatively comparable, yet quantitatively smaller response than capital goods production.

Moreover, the fact that the innovations to expectation cause long-run increases in manufacturing production in various sectors again points out that our expectation indicators effectively reflect information about future fundamentals. If the changes in firms' anticipations would mostly capture sentiments or noise, we would expect to find rather transitory effects on economic activity. Furthermore, the positive effect on impact, found for all except consumer non-durables, indicates that firms start to adjust right away before the higher level of productivity has actually been realized.

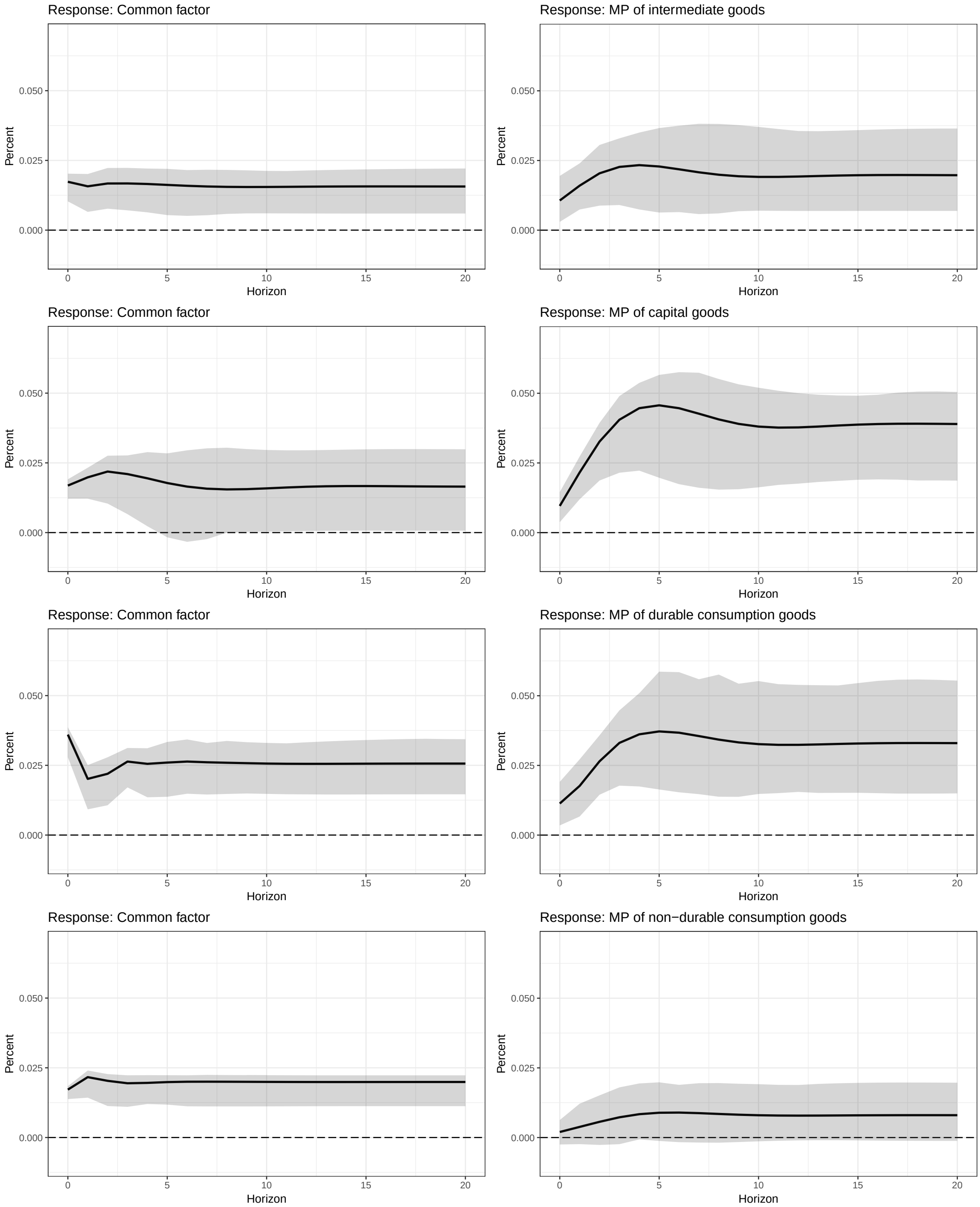

Figure 3.3: Cumulative impulse responses of manufacturing production (MP) across sectors to the common factor of the expectations (with 90 percent confidence intervals)

3.5 Conclusion

The business cycle literature has so far mainly focused on explaining and documenting aggregate comovement between macroeconomic aggregates such as investment and consumption. Yet, another crucial characteristics of business cycles is sectoral comovement, that is, the synchronized movements of inputs and outputs across sectors. Also the stylized fact of excess comovement, which means that production is stronger correlated across sectors than productivity, is regularly ignored in both theoretical and empirical works.

This paper examines the hypothesis of Veldkamp and Wolfers (2007) about the meaning of aggregate information for synchronized input and output movements across industries. Yet, the empirical support for this theory is still scarce as sector-level expectation data is hardly available. To fill this gap, we employ micro data from business tendency surveys to construct novel sector- and aggregate-level measures of anticipated productivity changes. Our indicator is found to track labour productivity which indicates that firms are able to anticipate its changes quite well.

The results of our factor analysis support the hypothesis of Veldkamp and Wolfers (2007) as we find that the firms in different sectors actually do base their expectations on similar information. Further, after controlling for the aggregate information component, the manufacturing production exhibits considerably lower correlations across disaggregated sectors. This indicates that common information plays a crucial role in explaining the synchronized sectoral movements.

Our further findings reveal that there is a great deal of heterogeneity in the reactions of the disaggregated manufacturing sectors to changes in expectations. In particular, expectation shocks are found to propagate mainly through the sectors producing capital goods and consumer durables. This is line with the expenditure side literature on news shocks (e.g., Beaudry et al. (2011) and Beaudry and Portier (2014)) as it theoretically argued and empirically found that investment reacts most to information about future fundamentals. Moreover, the finding that innovations to expectations measures cause long-run increases in manufacturing production in various sectors points to out that our expectation indicators effectively reflect information about future fundamentals.

Chapter 4

Is it good to be bad or bad to be good?: Assessing the impact of abnormal weather on consumer spending[1]

[1]This chapter is co-work with Boriss Siliverstovs. We thank MeteoSwiss for providing us with the access to the weather data.

4.1 Introduction

In the retail business, exceptional weather is often argued to have an impact on consumer spending. Such effects of abnormal weather typically manifest themselves as transitory shifts in consumer intertemporal spending decisions. Yet, formal support for this anecdotal evidence has hardly been found as the empirical literature on the impact of (abnormal) weather on consumer spending is still scarce. Obviously, usual weather effects can be easily quantified and removed by standard seasonal adjustment procedures. However, these procedures often struggle to appropriately accommodate the effects of weather anomalies, i.e., (very) untypical weather for a given season or month of the year, and may require non-conventional intervention in order to prevent distortions in the seasonal adjustment of the data of interest.

Understanding how unusual weather affects consumer spending is of importance for several reasons. On one hand, it reveals how consumer decisions are affected by abnormal weather conditions. Already Linden (1962) noted that unusual weather conditions cause shifts in timing of purchases, generate purchases that might otherwise not occur or cause a permanent loss of demand. Yet, the channels of exceptional weather on retail sales are multiple. First, weather may affect consumers' mood and therefore their spending decisions as argued by Murray et al. (2010). The more sunlight, the better the mood and the higher the willingness to spend (more) money. We refer to this as *mood channel*. Second, weather conditions also affect the convenience of the shopping experience (sunny weather vs., heavy rain or snow) and thus, increase or decrease, respectively, the motivation for shopping (*convenience channel*). Furthermore, weather conditions can boost sales of weather related products such as air conditioners, umbrellas and snow shovels (*weather-related products*). Moreover, when season changes there is need for different apparel as well as leisure equipment. Unusual bad or good weather can shift sales peaks during the months when new seasonal products are launched (*seasons change channel*). However, all these effects could be (partly) caught up in the following month(s). All in all, weather anomalies tend to affect the utility derived from a consumption good in a specific period and therefore cause intertemporal consumption patterns.

On the other hand, it is a well-known fact that consumer expenditure has a substantial transitory component that cannot be explained by changes in fundamental economic factors, such as income or interest rates. Unusual weather is assumed to be one of the main reasons for it. The transitory component can be problematic as it can distort the measurement of the effects of fundamental factors on spending. For instance, the estimation of elasticity of intertemporal substitution (EIS), one of the central parameters of macroeconomic and finance models, is affected by this issue. Furthermore, as retail sales is the most comprehensive measure of private expenditures at monthly frequency (national accounting data on private consumption is available on at quarterly frequency) and thus, is one of the main economic indicators, its movements are followed closely by central banks and economic analysts. Yet, it is crucial to disentangle if the observed changes are rather transitory (and possibly followed by a rebound in following month) or reflect genuine changes in underlying factors. Therefore, impacts caused by exceptional weather are relevant for business cycle analysis and monitoring current economic conditions as well as making projections in the future. Moreover, if abnormal weather affects retail sales, there could also be demand-led effects on inventories, production and employment, to mention but a few. In addition, the knowledge of weather impacts is also crucial for business planning and forecasting, especially for retailers.

Nevertheless, the effects of unusual weather conditions on consumer spending has so far received only limited attention in the empirical literature. The probably first study examining the transitory effects of weather on economic activity is the paper of Maunder (1973). Using weekly non-seasonally adjusted data, his findings indicate that weather conditions can account for a moderate share of the short-term variation in retail trade sales in the US. Also Starr-McCluer (2000) studies the effects of weather on (nominal) retail sales in the US on monthly and quarterly basis. She finds unusual hot and cold weather (measured by cooling and heating days) to have a significant but rather small effect on monthly nominal retail trades. Furthermore, her results reveal that the effects tend to differ depending on the quarter (the periodic analysis is done only for quarterly data). Yet, one problem with analysing unusual weather effects at national level in such a big country as the US is that weather tends to be very diverse across the country and therefore estimated impacts might average out at this level of aggregation. Busse et al. (2015) examine the influence of weather on car purchases in the US. They show that the choice of car type depends on the weather on the day of the purchase, even after controlling for the past weather conditions. Thus, their findings imply that consumers are affected by the weather conditions such that they tend to overstate the future utility when buying high-value durable goods.

The main aim of our analysis is to investigate if weather anomalies lead to intertemporal shifts in consumer spending at monthly frequency and to quantify the size of these effects. We contribute to the existing literature in different ways. In relation to Maunder (1973) who considers only contemporaneous effects and uses only three years of data, we analyse longer time series and allow also for rebound impacts to get more reliable evidence. Compared to Starr-McCluer (2000) we use more precise weather measures using actual temperature instead of number of cooling or heating days only. More importantly, we conduct the periodic analysis on monthly level to be able to discover the exact nature of abnormal weather effects which tend to "wash out at a quarterly frequency" as noted by Starr-McCluer (2000). Murray et al. (2010), like most of the marketing literature, employ data of a particular store and specific product groups. Nevertheless, for statements of total (nation wide) impacts, aggregate sales data is required. Oppose to Busse et al. (2015) who concentrate on very specific market and do not use nation-wide data, we analyse various product groups containing aggregate retail sales data, which allows us to make conclusions about the total impacts of abnormal weather on consumer expenditures. In particular, we add on all of these works offering theoretical model based on consumer choice. Moreover, the application of long-run restrictions and quantification of these effects borrowing the approaches of Boldin and Wright (2015) who examine the weather-adjustment of employment data, is new in weather-related consumer spending literature.

In Section II, we develop a stylized model of intertemporal consumption in which abnormal weather affects consumption enjoyment, i.e., the utility derived from consumption. These effects are allowed to vary across periods as the impact of weather anomalies is very likely to differ depending on the month.

In Section III we then use Swiss retail sales and weather data as described to test if exceptional weather does influence consumer decisions as proposed by the theoretical model. As Switzerland is a small country where weather does not differ greatly between the highly populated areas (where most retail sales are made), we can conduct this analysis using national macroeconomic data and thus, be able to evaluate how large the impacts of abnormal weather at the aggregate level are. We consider three different weather variables: temperature, precipitation and sunshine. In line with the literature, the abnormal weather measures are defined as deviations from their month-specific long-term (running) mean.

Section IV documents our empirical results. Our findings reveal that weather anomalies do cause intertemporal shifts in consumer expenditure in Switzerland. They can explain a considerable share of the variability of seasonally-adjusted retail sales, especially in the non-food sector. We find that consumers react at most to exceptional temperatures and less to exceptional precipitation or sunshine, implying that temperature is the most influential weather variable for explaining the intertemporal shifts in consumer spending. Furthermore, the effects of abnormal weather are found to differ across seasons, i.e., to be month-specific, both with respect to the sign and to the magnitude. In particular, our findings indicate that weather effects manifest mainly through the seasons change channel: exceptionally warm temperatures in early spring (good to be good) as well as unusually cold conditions in late summer and early autumn (good to be bad) are generally associated with higher sales than usual. That is, weather conditions in line with the coming season motivate to make the purchases early in the season. In other words, depending on the season (or month) unusually good weather may boost or restrain consumer expenditures, and vice versa.

Section V presents the results of several extensions of our analysis presented in the main body of the paper aiming to verify their robustness. The final section concludes.

4.2 Theoretical model

To formalize our arguments theoretically, we develop a stylized model of intertemporal consumption. In this model, weather conditions affect consumption enjoyment i.e., the utility derived from consumption. As discussed above, this can be due to the mood, convenience, weather related products and seasons change effects. Furthermore, we allow the importance of these weather conditions to vary seasonally. Assuming an isoelastic utility function, this can be formalized as follows

$$U_t = \frac{C_t^{1-\gamma}}{1-\gamma} s_t^{\theta_m}, \tag{4.1}$$

where s_t stands for the weather state taking value 1 by average (normal) weather, θ_m indexes the importance of the weather state and $m = 1, \ldots, 12$ denotes the month in which t falls. s_t can be interpreted as a *taste-shifter*, a variable that shifts marginal utility. We assume life time utility to be additive so that

$$V_t = \sum_{i=0}^{T} \beta^i U_{t+i}, \quad 0 < \beta \leq 1, \tag{4.2}$$

where β is a time discount factor. The budget constraint is defined as

$$C_{t+1} = Y_{t+1} - A_{t+1} + (1 + r_{t+1})A_t, \tag{4.3}$$

where Y is real income, r is the real interest rate and A is the end-of-period real value assets. Maximization of total utility yields

$$\frac{C_{t+1}^{-\gamma} s_{t+1}^{\theta_{m+1}}}{C_t^{-\gamma} s_t^{\theta_m}} = \frac{1}{\beta(1 + r_{t+1})}. \tag{4.4}$$

Taking logarithms and adding the disturbance gives us

$$\gamma \ln C_{t+1} + \ln s_{t+1}^{\theta_{m+1}} = \gamma \ln C_t + \ln s_t^{\theta_m} - \ln\left(\frac{1}{\beta(1+r_{t+1})} + \epsilon_{t+1}\right). \tag{4.5}$$

We assume that the stochastic term is normally and identically distributed

$$\ln\left(\frac{1}{\beta(1+r_{t+1})} + \epsilon_{t+1}\right) \sim N(\mu, \sigma^2).$$

Thus, we can use the properties of log-normal distributions to derive the following results:

$$E_t\left(\frac{1}{\beta(1+r_{t+1})} + \epsilon_{t+1}\right) = \exp(\mu + 1/2\sigma^2)$$

and further,

$$\mu = \ln(1/\beta(1+r_{t+1})) - 1/2\sigma^2.$$

Finally, we can write the equation (5) as

$$\ln C_{t+1} = \omega + \frac{\ln(1+r_{t+1})}{\gamma} + \ln C_t + \theta_m \ln s_t - \theta_{m+1} \ln s_{t+1} + u_{t+1}, \tag{4.6}$$

where

$$\omega_t = \frac{1}{\gamma}\ln(\beta + 1/2\sigma^2).$$

Reordering the terms gives us the final specification:

$$\ln\frac{C_{t+1}}{C_t} = \omega + \frac{\ln(1+r_{t+1})}{\gamma} + \theta_m \ln s_t - \theta_{m+1} \ln s_{t+1} + u_{t+1}. \tag{4.7}$$

Equation (4.7) states that the growth in consumption depends on the time discount factor (β), the interest rate, the weather state of the current period as well as the weather state of the previous period and the forecast error. This implies that through intertemporal optimization unusual weather may cause shifts in consumption over time.

Often it is assumed that the utility does not depend only on the current consumption expenditure but also on the previous consumption level. A utility function with habit formation can be written as

$$U_t = \frac{1}{1-\gamma}\tilde{C_t}^{1-\gamma} s_t^{\theta_m} \tag{4.8}$$

where $\tilde{C_t} = \frac{C_t}{C_{t-1}^{\phi}}$ and ϕ controls the importance of habit formation. Then, the maximization of total utility yields

$$\left(\frac{C_{t+1}}{C_t^{\phi}}\right)^{-\gamma} \frac{s_{t+1}^{\theta_{m+1}}}{C_t^{\phi}} = \left(\frac{C_t}{C_{t-1}^{\phi}}\right)^{-\gamma} \frac{s_t^{\theta_m}}{C_{t-1}^{\phi}} \frac{1}{\beta(1+r_{t+1})}. \tag{4.9}$$

Under the same assumptions as earlier, we get

$$\ln\frac{C_{t+1}}{C_t} = \omega + \frac{\ln(1+r_{t+1})}{\gamma} + \theta_m \ln s_t + \theta_{m+1}\ln s_{t+1} + (1-\frac{1}{\gamma})\phi\ln\frac{C_t}{C_{t-1}}. \tag{4.10}$$

Now, the growth of consumption depends also on the growth rate of the previous period.

4.3 Data

For this analysis we employ three data sets: data on weather, retail sales and macroeconomic variables such as interest rates and inflation. The weather data for this paper comes from the Swiss Federal Office of Meteorology and Climatology (MeteoSwiss)[2]. The various weather variables are available for numerous weather stations in Switzerland. The national values are defined as simple average of 12 specific weather stations[3]. Since we want to examine the effects of unusual weather, we construct our weather variables as deviations from the month-specific long-run

[2]For the data access, see https://gate.meteoswiss.ch/idaweb/

[3]Stations: BAS (Basel), BER (Bern), CHD (Château-d'Oex), CHM (Chaumont), DAV (Davos), ENG (Engelberg), GVE (Geneva-Cointrin), LUG (Lugano), SAE (Saentis), SIA (Segl-Maria), SIO (Sion), SMA (Zurich)

mean following international standards[4]. First, we calculate the national averages of 12 stations for the monthly values. Then the rolling 30-years mean for each month is computed. Finally, we define the deviations as the monthly value minus the (one-year) lagged rolling mean as follows:

$$W_t^m = w_t^m - \frac{1}{30} \sum_{t-31}^{t-1} w_t^m \tag{4.11}$$

where w_t^m denotes the value of the weather variable in month m. We repeat this for all the three weather variables we consider: homogenized monthly mean temperature (2 meters above the ground), homogenized monthly mean precipitation (in millimeters) and monthly mean duration of sunshine (in hours). The sample period of the weather variables runs from January 1980 to February 2017. These variables are plotted in Figure 4.1.

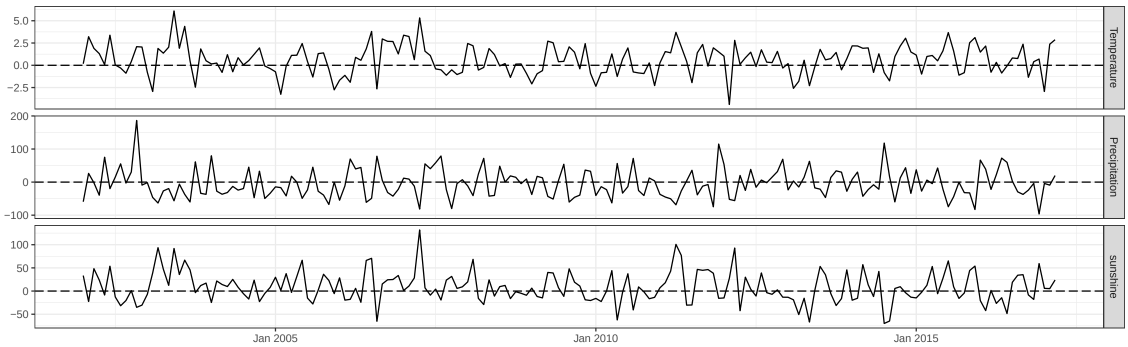

Figure 4.1: Weather variables - deviations from long-run rolling mean

The Swiss Federal Statistical Office (FSO) provides data on retail trade sales in Switzerland. The indexes start on January 2002. The data are available for total retail trade (NOGA 47) and for the sub-branches. Our main series are the seasonally and calender effects adjusted retail sector without fuel (NOGA 47 without NOGA 473)[Total wo fuel], retail sales of food, beverages and tobacco (NOGA 4711 and

[4]See, Definition of World Meteorological Organization to Climatological Normals: http://www.wmo.int/pages/prog/wcp/wcdmp/GCDS$_1$.*php*

472) [Food] as well as retail sales of non-food (NOGA 4719, 474-479) [Non-Food]. Figure 4.2 shows the series over the sample period from January 2002 to December 2016.

Moreover, the data on short-term nominal interest rates as well as CPI index we extract from the SNB Dataportal for the same sample period (January 2002 - December 2016).

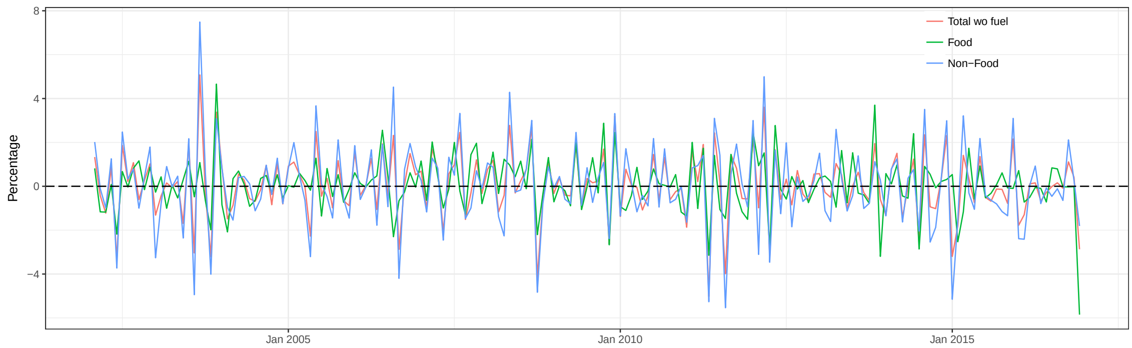

Figure 4.2: Nominal Retail sales (Month-to-Month growth rates)

4.4 Empirical analysis

In this section, we present our estimation results for several model specifications. In Section 4.1, we estimate our baseline model where we impose identical weather effects for all months. In Section 4.2, we relax this assumption and allow for month-specific or periodic weather effects. In Section 4.3, we test the null hypothesis that there are no long-run effects of abnormal weather on the level of consumption.

4.4.1 Baseline specification

To examine the short-term effects of unusual weather on retail trade sales, we estimate various regression specifications derived from equation (4.7) of our theoretical model.

In the first step, we consider constant weather effects over the year i.e., assume that θ do not vary seasonally:

$$\Delta C_t = \omega + \sum_{i=1}^{p} a_i \Delta C_{t-i} + \delta R_t + \sum_{l=0}^{1} \theta_l W_{t-l} + \epsilon_t, \tag{4.12}$$

where C_t is log nominal retail sales (index) at time t, p indicates the number of autoregressive lags, $R_t = \ln(1+r_t)$ whereas r_t is the short-term interest rate and $\delta = 1/\gamma$ captures the elasticity of intertemporal substitution. $W_{t-l} = \{\text{Temp, Rain, Sun}\}$ is one of our weather variables and l denotes a lag.

Coefficient estimates for this equation are reported in Table 4.1. In column (1), we do not include any weather variables so that we can later compare the results with this baseline specification. In column (2), we include the contemporaneous and lagged values of the temperature variable, in the third and fourth columns - those of the precipitation and the sunshine. In column (5), the estimation result of a model with all weather variables is presented.

Table 4.1: Average weather effects

	Dependent variable:				
	Nominal Retail turnover (dl)				
	(1)	(2)	(3)	(4)	(5)
Constant	0.041 (0.093)	0.050 (0.101)	0.036 (0.094)	0.027 (0.097)	0.037 (0.102)
Δc_{t-1}	-0.756^{***} (0.075)	-0.748^{***} (0.075)	-0.752^{***} (0.075)	-0.756^{***} (0.075)	-0.738^{***} (0.076)
Δc_{t-2}	-0.534^{***} (0.085)	-0.521^{***} (0.085)	-0.531^{***} (0.085)	-0.532^{***} (0.086)	-0.505^{***} (0.086)
Δc_{t-3}	-0.217^{***} (0.075)	-0.209^{***} (0.075)	-0.214^{***} (0.075)	-0.215^{***} (0.075)	-0.198^{***} (0.075)
R_t	0.452^{***} (0.119)	0.446^{***} (0.119)	0.454^{***} (0.120)	0.448^{***} (0.120)	0.425^{***} (0.122)
Temp_t		-0.081 (0.056)			-0.144^{**} (0.070)
Temp_{t-1}		0.066 (0.056)			0.082 (0.070)
Rain_t			-0.001 (0.002)		-0.0003 (0.003)
Rain_{t-1}			0.0002 (0.002)		0.001 (0.003)
Sun_t				0.001 (0.003)	0.005 (0.005)
Sun_{t-1}				0.001 (0.003)	0.001 (0.005)
Observations	176	176	176	176	176
R^2	0.389	0.399	0.390	0.390	0.409
Adjusted R^2	0.374	0.378	0.368	0.368	0.373

Note: s.e. in parentheses $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Our estimation results suggest that when restricting the weather effects to be equal for all months of the year, there is a very limited evidence suggesting that abnormal weather influences retail sales in Switzerland. Almost all but one ($Temp_t$ in column (5)) weather variables are insignificant and the adjusted R^2 increases only marginally for models in columns (2) and (5) out of all models augmented with weather variables (columns (2)-(5)) as compared with our benchmark model in column (1) that excludes all weather variables.

4.4.2 Month-specific weather effects

Next, we examine if consumers are affected by weather anomalies differently at different seasons or months. Therefore, we allow the weather effects to differ periodically, i.e., θ may vary over the year:

$$\Delta C_t = \omega + \sum_{i=1}^{p} a_i \Delta C_{t-i} + \delta R_t + \sum_{l=0}^{1} \theta_{1,l} D^{Jan} W_{t-l} + \sum_{l=0}^{1} \theta_{2,l} D^{Feb} W_{t-l} + \cdots + \sum_{l=0}^{1} \theta_{12,l} D^{Dec} W_{t-l} + \epsilon_t \quad (4.13)$$

where $W_{t-l} = \{\text{Temp, Rain, Sun}\}$ is one of our weather variables and D^m a dummy variable taking value 1 in month m and zero otherwise.

The results for the temperature variable are presented in Table 4.2. Again, in the first column the estimated coefficients for the baseline model without weather variables are documented. The second column presents the results of the model with month-specific temperature effects. Numerous temperature coefficients turn out to be statistically significant, however, the signs and the sizes of the coefficients differ greatly. The estimate for March is found to be positive indicating that unusual warm weather in March boost the retail sales. The coefficient on June temperature is, in turn, found to be negative implying that abnormal hot weather during the summer month exercises a dampening effect on the retail sales. From August to October the coefficients on the temperature dummies are negative and significant, with the October coefficient having the highest value. This finding implies that abnormal warm weather conditions during early autumn have negative impact on the retail sales hindering changes in the garderope and shifting the seasonal sales peaks later.

We repeat the same analysis using precipitation and sunshine as weather variables. The results are documented in Tables B.1 and B.2 in B Appendix. Nevertheless, the results imply similar impacts as the model with temperature: abnormal nice weather (less rain or more sunshine, respectively) in spring months boost the sales of retailers, whereas it hinders the sales in early autumn. However, hardly any of the coefficients are significant.

Table 4.2: Month-specific effects of Temperature

	Dependent variable:	
	Nominal Retail turnover (dl)	
	(1)	(2)
Constant	0.041 (0.093)	0.033 (0.097)
Δc_{t-1}	-0.756^{***} (0.075)	-0.688^{***} (0.077)
Δc_{t-2}	-0.534^{***} (0.085)	-0.470^{***} (0.084)
Δc_{t-3}	-0.217^{***} (0.075)	-0.164^{**} (0.071)
R_t	0.452^{***} (0.119)	0.509^{***} (0.115)
$D^{Jan}Temp_t$		0.074 (0.187)
$D^{Feb}Temp_t$		0.028 (0.140)
$D^{Mar}Temp_t$		0.503^{**} (0.199)
$D^{Apr}Temp_t$		0.213 (0.139)
$D^{May}Temp_t$		0.068 (0.209)
$D^{Jun}Temp_t$		-0.315^{**} (0.145)
$D^{Jul}Temp_t$		0.197 (0.176)
$D^{Aug}Temp_t$		-0.615^{***} (0.156)
$D^{Sep}Temp_t$		-0.327^{*} (0.182)
$D^{Oct}Temp_t$		-0.697^{***} (0.214)
$D^{Nov}Temp_t$		-0.241 (0.162)
$D^{Dec}Temp_t$		0.195 (0.210)
$D^{Jan}Temp_{t-1}$		0.114 (0.202)
$D^{Feb}Temp_{t-1}$		-0.152 (0.192)
$D^{Mar}Temp_{t-1}$		-0.160 (0.119)
$D^{Apr}Temp_{t-1}$		-0.093 (0.206)
$D^{May}Temp_{t-1}$		-0.374^{**} (0.153)
$D^{Jun}Temp_{t-1}$		0.187 (0.207)
$D^{Jul}Temp_{t-1}$		-0.049 (0.146)
$D^{Aug}Temp_{t-1}$		0.271^{*} (0.157)
$D^{Sep}Temp_{t-1}$		0.461^{***} (0.163)
$D^{Oct}Temp_{t-1}$		0.416^{**} (0.198)
$D^{Nov}Temp_{t-1}$		0.157 (0.202)
$D^{Dec}Temp_{t-1}$		0.399^{**} (0.196)
Observations	176	176
R^2	0.389	0.608
Adjusted R^2	0.374	0.534
Residual Std. Error	1.190 (df = 171)	1.030 (df = 147)
F Statistic	27.200^{***} (df = 4; 171)	8.150^{***} (df = 28; 147)
Note: s.e. in parentheses		*p<0.1; **p<0.05; ***p<0.01

Various coefficients on the lagged weather variables are also found to be significant. Again, depending on the month, the sign of the estimates differ. The coefficient $Temp_{t-1}^{May}$ tells us that the unusually warm weather in April has a negative effect on retail sales in May. On the other hand, the abnormally high temperatures in August, imply a positive rebound effect in September.

The estimated coefficients for temperature are also graphically illustrated in Figure 4.3 showing again temperature effects being positive in the first half of the year and become in general negative in the second half of the year. For the previous month's effect the picture is basically the opposite, the rebound effects being negative in first five months and turning positive from August on which is consistent with the rebound concept.

Our results allow us to draw several conclusions regarding the relative importance of different channels through which abnormal weather is expected to affect consumer spending decisions. These findings do not support the mood channel since unusually good weather (warm, sunny, less rain) is found to have in some months positive and other months negative effects, if the mood effect would be the main channel we would expect the coefficients be always positive. Also the convenience channel does not find great support for similar reasons, bad weather seems to boost the sales in specific months. The weather-related products do not seem to play crucial role either since the impacts in winter or summer months are quantitatively small and not significant, expect one. Yet, we find strong support for the seasons change effects. Abnormal high temperatures foster seasonal product sales in spring i.e., make new seasonal products more appealing. On the other hand, especially cold weather conditions lead to higher sales in late summer and early autumn. That is, weather conditions in line with the coming season motivate to make the purchases early in the season. All in all, this implies that in some months warm weather boost consumer spending, whereas in other months cold temperatures induce more sales.

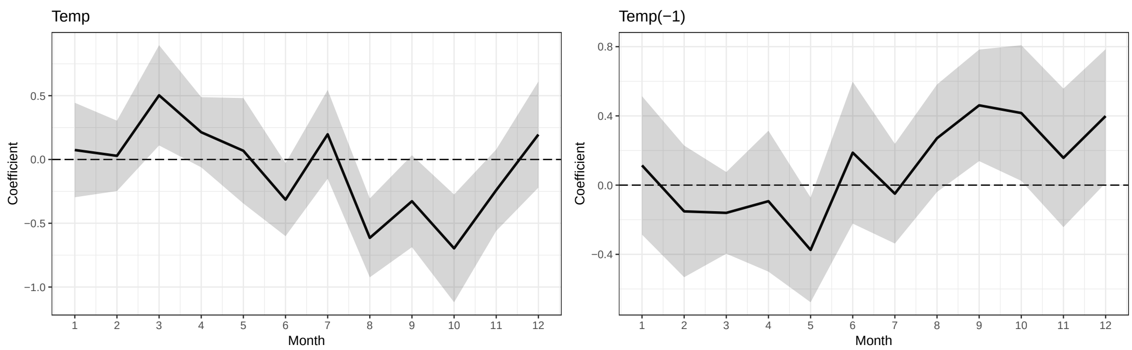

Figure 4.3: Month-specific coefficients of temperature variables

All in all, abnormal weather is found to be able to explain considerable share of the variance of retail sales growth since the adjusted R^2 increases considerably. The other weather variables (precipitation and sunshine) seems to matter less as the adjusted R^2 hardly increases. Our findings suggest that there are positive effects of abnormal weather on the retail sales that can be realized equally well for unusually hot or cold days, depending on a season, or vice versa. Moreover, these effects are further reinforced because of rebound effects, i.e., what is over- or under-consumed in a given month tends to be caught up in the following month.

Moreover, since we might have problem overfitting as the number of parameters is high compared to the number of observations, we check the robustness of the results in Table 4.2 applying variable selection based on LASSO analysis. In Table B.3 in B Appendix we report the results of the re-estimated model with fewer variables as suggested by LASSO procedure. The main findings still remain the same.

4.4.3 Long-run restrictions

In the previous section, we documented a rebound effect of the abnormal weather on retail sales. In this section, we test whether over- or under-consumption that take place in a given month tends to be exactly compensated in the following month such that there are no long-run effects brought by unusual weather on retail trade. In doing so, we follow Boldin and Wright (2015) (BW) and define monthly (dummy) variables such that they take 1 in a specific month (to capture the current weather effect) and -1 in the following month but for previous month's weather, i.e., equal size but opposite sign weather effect is imposed for the following month, and 0 otherwise. This implies that weather shocks cannot have permanent effects on the level of retail sales but will be followed by a bounce back in the following month or later through the autoregressive dynamics. One further advantage of testing and eventually correctly imposing these restrictions is that it allows us significantly reduce the number of parameters to be estimated in our regressions, thus mitigating the potential problem of overfitting.

The estimation results for total retail sales (without fuel), food and non-food sector are shown in Table 4.3. The likelihood ratio test indicates that we cannot reject these BW-restrictions, i.e., there is no evidence for permanent weather effects. The findings in column (1) are similar to the earlier results that is we find a positive effect of excess temperatures in the early spring months, adverse effect in June and from August to November. In the other two columns we document the results for food and non-food sectors separately. The findings suggest that weather affects mainly the non-food sales, since the estimation results for non-food are similar but stronger than for total retail sales without fuel. For food, the temperature effects are not found to be significant in the first half of the year, but coefficients on weather variables from October to December are negative and significant, however smaller size than for non-food.

Table 4.3: Month-specific effects of Temperature with long-run restriction

	Dependent variable:		
	Total	Food	Non-Food
	(1)	(2)	(3)
Constant	0.035 (0.080)	0.079 (0.089)	−0.024 (0.102)
Δc_{t-1}	-0.602^{***} (0.069)		
Δc_{t-2}	-0.404^{***} (0.078)		
Δc_{t-3}	-0.123^{*} (0.068)		
Δc_{t-1}		-0.503^{***} (0.082)	
Δc_{t-2}		-0.263^{***} (0.090)	
Δc_{t-3}		−0.095 (0.084)	
Δc_{t-1}			-0.567^{***} (0.068)
Δc_{t-2}			-0.434^{***} (0.074)
Δc_{t-3}			-0.161^{**} (0.066)
R_t	0.390^{***} (0.104)	0.330^{***} (0.113)	0.454^{***} (0.131)
$D^{Jan}Temp_t$	0.181 (0.124)	−0.027 (0.143)	0.346^{**} (0.156)
$D^{Feb}Temp_t$	0.105 (0.088)	−0.018 (0.098)	0.211^{*} (0.111)
$D^{Mar}Temp_t$	0.353^{**} (0.143)	0.152 (0.158)	0.476^{***} (0.182)
$D^{Apr}Temp_t$	0.301^{***} (0.097)	0.170 (0.106)	0.272^{**} (0.122)
$D^{May}Temp_t$	−0.045 (0.137)	−0.022 (0.152)	−0.122 (0.174)
$D^{Jun}Temp_t$	-0.167^{*} (0.094)	0.014 (0.103)	-0.252^{**} (0.119)
$D^{Jul}Temp_t$	−0.094 (0.113)	0.180 (0.125)	−0.218 (0.143)
$D^{Aug}Temp_t$	-0.563^{***} (0.114)	−0.068 (0.121)	-0.833^{***} (0.147)
$D^{Sep}Temp_t$	-0.337^{**} (0.131)	−0.101 (0.144)	-0.549^{***} (0.166)
$D^{Oct}Temp_t$	-0.423^{***} (0.143)	-0.296^{*} (0.151)	-0.503^{***} (0.186)
$D^{Nov}Temp_t$	-0.375^{***} (0.116)	-0.301^{**} (0.129)	-0.419^{***} (0.145)
$D^{Dec}Temp_t$	0.093 (0.131)	-0.292^{**} (0.145)	0.293^{*} (0.167)
LR-test (p-value)	0.229	0.693	0.304
Observations	176	176	176
R^2	0.573	0.294	0.603
Adjusted R^2	0.530	0.223	0.563
Residual Std. Error (df = 159)	1.030	1.140	1.310
F Statistic (df = 16; 159)	13.300^{***}	4.150^{***}	15.100^{***}

Note: s.e. in parentheses $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

For some months the sign of the weather impact is the different for food and non-food sales indicating opposite effects. Yet, only for December, the coefficients for both product groups are significant. The unusually warm weather boost the non-food sales whereas it found have a negative effect on food sales.

Unlike Busse et al. (2015), we do not find evidence for projection bias since we find that unusual temperatures lead mainly to shifts in the purchase timing. Yet, we consider also very different product groups. Busse et al. (2015) examine car sales, that is, highly durable and specialized products, whereas we examine the retail sales (which do not include car sales) but include also many non-durable categories. It seems that in our case the abnormal temperatures supporting the coming season trigger consumer to make their purchases early in the season.

Table 4.4: Incremental adjusted R^2

	Temperature
Total	0.1554
Food	0.0219
Non-Food	0.1725

Also the increase in adjusted R^2 is the highest for the non-food retailers as shown in Table 4.4 implying that the unusual temperatures can capture considerable share of the volatility of retail sales. Altogether, abnormal weather seems to cause inter-temporal effects as consumers do shift their non-food purchases depending on the temperature conditions.

To assess the quantitative meaning of these weather impacts, we firstly calculate the partial R^2 for the month-specific weather variables in order to see which months are especially meaningful. The bars in Figure 4.4 tells us the proportion of variation explained by each month that cannot be explained by the other variables. The highest value is found for August indicating that the temperature in August can explain around 13 % of the variation in the total retail sales, for non-food even more than 15 % cannot be captured by the other variables.

Another way to quantify the weather impact is to define counterfactual series as in Boldin and Wright (2015). The counterfactual series are calculated by setting the weather indicators to zero but using the same residuals. The difference between the original series and the counterfactual series can then be interpreted as the weather effect. We make these calculations for the models in Table 4.3. As shown in Figure 4.5 the highest median absolute weather effects are found in August, September and November of size around 0.6-0.7 percentage points indicating that temperature anomalies can account for a noticeable change in the retail sales growth. The biggest

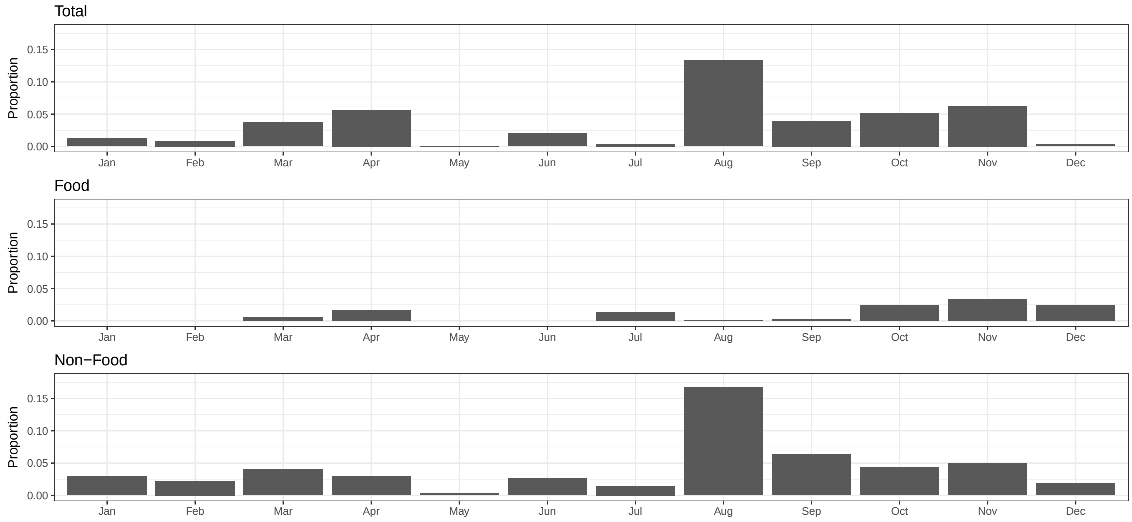

Figure 4.4: Partial R-squared
Note: Upper panel: Model in column (1) in Table 3, middle panel: column(2) in Table 3, lower panel: column (3) in Table 3

(absolute) contribution of unusual temperature in September is as high as 2.5 percentage points. For almost half of the months is the median absolute temperature effect over 0.5 percentage points.

Again, it is obvious in Figure 4.5 that the non-food sector is influenced much more strongly by exceptional temperatures than the food retail sales. Here, the highest median effect is found for August and it counts for more than 1 percentage point, whereas the greatest impact of almost 4 percentage points is found for September. Altogether these findings imply that the influence of abnormal temperature on retail trade is sizeable.

4.5 Robustness

To examine if our results are sensitive to the definition of the dependent variable, we estimate also the main equations using real retail sales instead of nominal as well as year-to-year growth rates instead of month-to-month specification. Table B.4 in B Appendix provides the results for the specification in real terms. The outcomes are very much like those in Table 4.3 in B Appendix. We again found the expected negative sign for the real short-term interest rates.

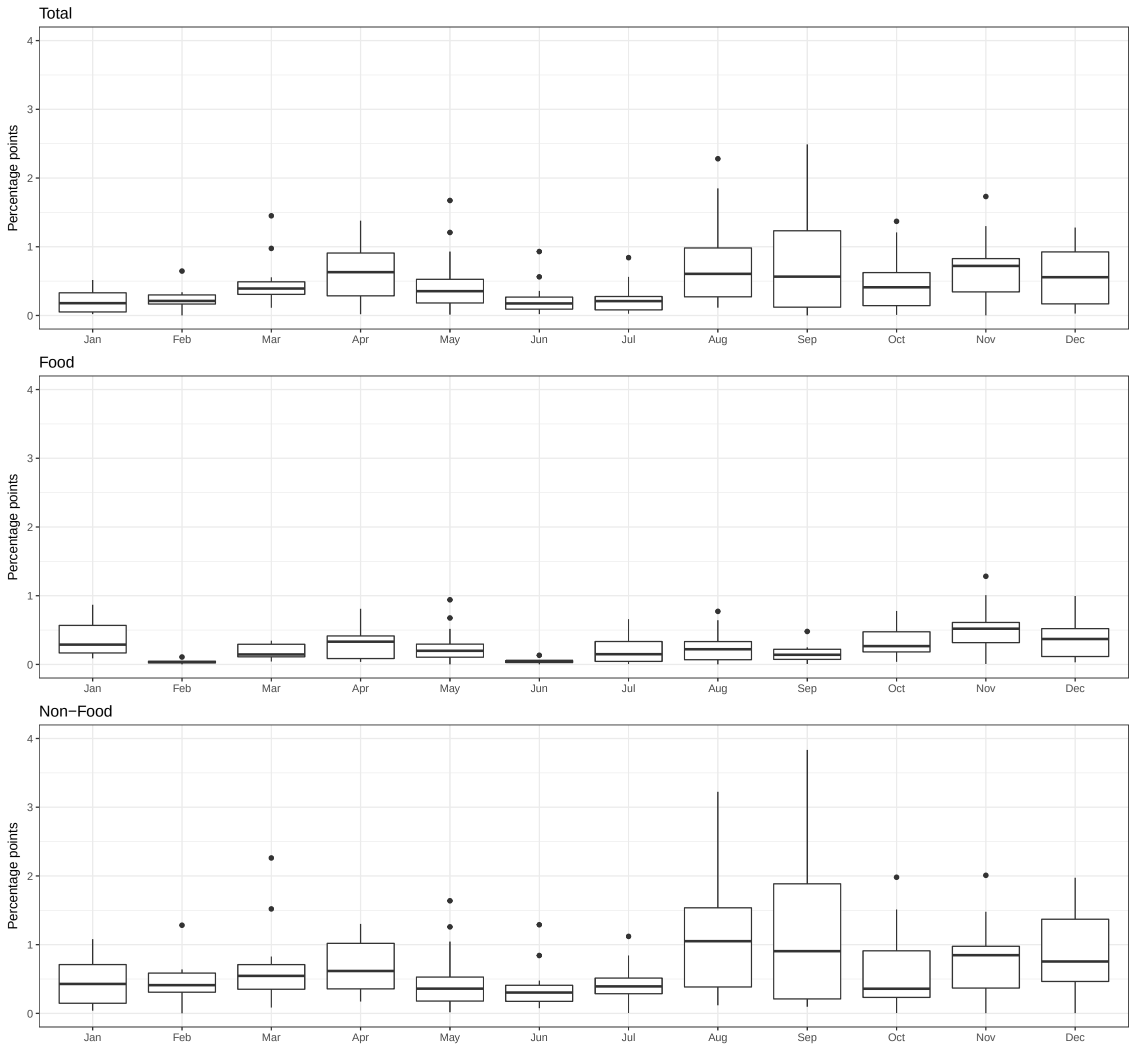

Figure 4.5: Absolute month-specific temperature effects
Note: Upper panel: Model in column (1) in Table 3, middle panel: column(2) in Table 3, lower panel: column (3) in Table 3)

The impact of abnormal temperature is also found to be quite similar when the year-on-year growth rates of nominal retail sales are used instead of month-to-month changes (Table B.5 in B Appendix). In the third robustness check we control for unemployment rate to make sure that the results are not sensitive for labour market situation. The results are reported in Table B.6 in B Appendix. In column (1) the change in unemployment rate whereas in column (2) the level of unemployment rate is used. In both cases the findings concerning the weather effects do not change. Further, we also examine if including exchange rate affects the results. This is not the case as shown in Table B.7 in B Appendix.

4.6 Conclusions

Already Linden (1962) noted that unusual weather conditions cause shifts in timing of purchases, generate purchases that might otherwise not occur or cause a permanent loss of demand. Also in the business press, exceptional weather is often argued to have an impact on consumer spending and business activity in general, being one of the main causes for the transitory shifts.

We contribute to the so far scarce literature by a comprehensive periodic analysis at monthly frequency accounting for both contemporaneous as well as lagged effects and using precise weather measures. Moreover, the application of long-run restrictions and quantification of these effects borrowing the approaches of Boldin and Wright (2015) who examine the weather-adjustment of employment data, is new in weather-related consumer spending literature. In addition, we develop a stylized model based on consumer choice in order to illustrate how abnormal weather affects the utility and therefore, leading into intertemporal shifts in consumption.

Our findings reveal that weather anomalies do cause substantial intertemporal shifts - the long-run restrictions cannot be rejected - in consumer expenditure in Switzerland. They can explain a considerable share of the variability of seasonally-adjusted retail sales, especially in the non-food sector. We find that consumers react at most to exceptional temperatures, less to abnormal precipitation or sunshine, implying that temperature is the most influential weather variable for explaining volatility of retail sales. Furthermore, the effects of abnormal weather are found to differ across seasons i.e., to be month-specific, both the sign and the magnitude. In particular, our findings indicate that weather effects manifest mainly through the seasons change channel: exceptionally warm weather in spring tends to boost the sales (good to be good), whereas unusually cold conditions in late summer/early autumn are generally associated with higher sales (good to be bad). That is, weather conditions in line with the coming season motivate to make the purchases early in the season. In other words, depending on the season (or month) unusually good or bad weather may boost or restrain consumer expenditures.

Chapter 5

Detecting outliers in small samples from skewed and heavy-tailed distributions

5.1 Introduction

Outliers are usually defined as observations which appear to be incoherent with the majority of observations of a data set (cf. Barnett and Lewis (1994)). As already Barnett and Lewis (1994) note, outliers are a subjective concept (*appear to be incoherent*) as well as a relative concept (*the majority of observations of a data set* or the underlying population). Closely related notation is a *contaminant* which refers to an observation from a different population than the main population. However, not all contaminants are suspect, for example, if they occur in the middle of the data set. Only when they are also *extreme observations*, that is, the smallest or the largest data points of a sample, they might be outlying. Yet, not even all extreme observations are outliers, only if they are *far enough* from the other data points, they are potentially considered as outliers. To identify the unusual observations in practice, many outlier detection methods have been proposed. They aim to provide guidance on which data points should be considered as outlying.

Further, Chambers (1986) was first to differentiate between *representative* and *non-representative* outliers. The former are observations with correct values and are not considered to be unique, whereas non-representative outliers are elements with incorrect values or are for some other reason considered to be unique. Most of the outlier analyses focus on the representative outliers as non-representatives values should be taken care of already in data editing and validation.

The main reason to be concerned about possible outliers is that calculated estimates might differ considerably depending on whether or not they are included in the sample. The data points with substantial influence on the estimates are also called *influential observations* and they should be identified and if needed, treated in order to ensure unbiased results.

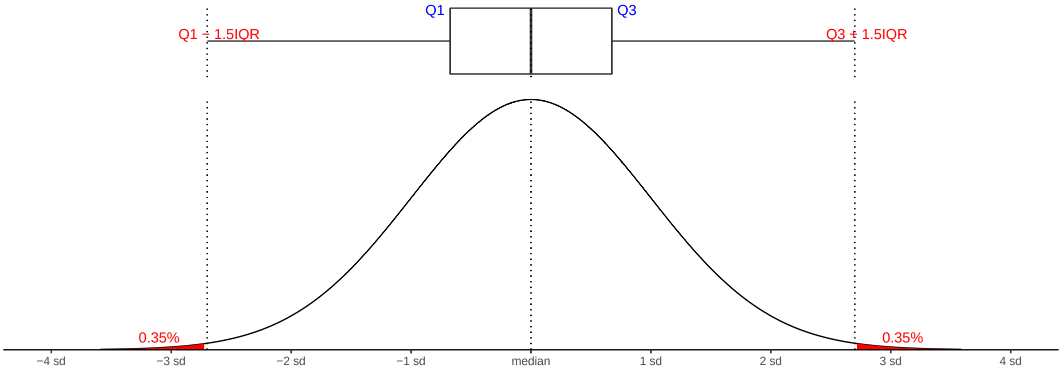

Figure 5.1: Standard boxplot for standard normal distribution

Owing to the fact that most of the outlier detection methods assume that the underlying population is normally distributed, in case of non-normal distribution too many or the wrong observations are potentially declared as outlying. Thus, if the base distribution itself is assumed to be asymmetric or heavy-tailed, this needs to be taken into account in outlier analysis. For example, resistant rules, also known as boxplot, which was first proposed by Tukey (1977), are very popular in outlier labelling (displayed in Figure 5.1). The problem with skewed data is, however, that one tends to identify (too) many outliers in the long tail but hardly any in the short tail as illustrated in Figure 5.2. Yet, in the case of a symmetric, but heavy-tailed sample, numerous observations tend to be outside the fences on both sides as shown in Figure 5.3. The same issues also concern other outlier detection techniques such as relative distance based methods.

Even though a large number of techniques to identify outlying observations have already been proposed (see, for example Barnett and Lewis (1994)), methods to deal with outliers in skewed or heavy-tailed data are still scarce. Hubert and Vandervieren (2008) propose an adjusted boxplot for skewed distribution in order to address the issue that in the case of skewed data the standard boxplot often labels incorrectly too many observations in the upper tail as outliers. Bruffaerts et al. (2014) go even further and suggest a generalized boxplot for skewed and heavy-tailed distributions based on the *Tukey g-and-h distribution*. A similar idea is behind the work of Xu et al. (2014) who present an outlier detection approach relying to the Tukey's g-and-h distribution.

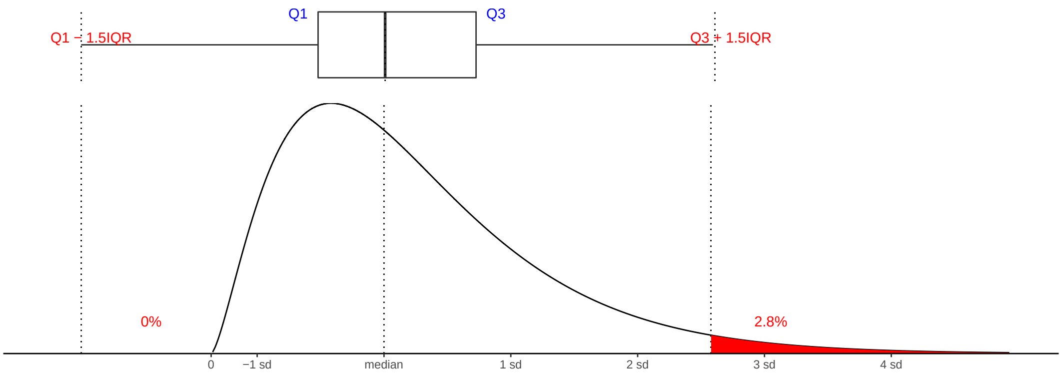

Figure 5.2: Standard boxplot for χ^2_5-distribution

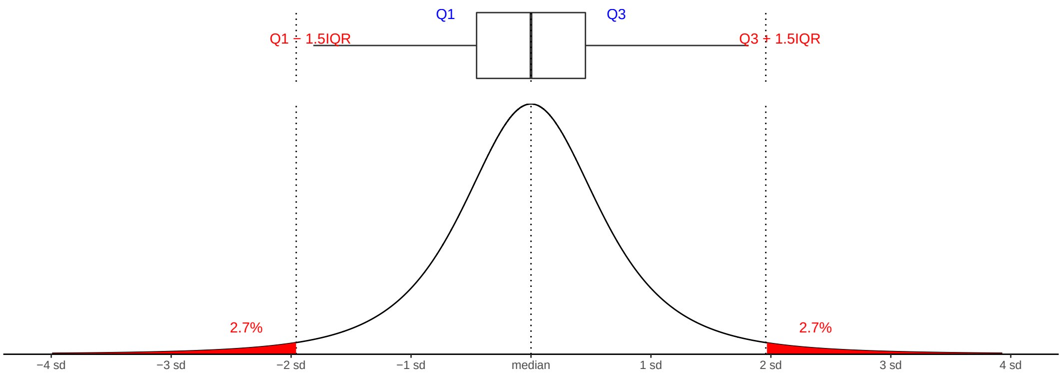

Figure 5.3: Standard boxplot for $t(3)$-distribution

In general, outlier detection methods are derived asymptotically. However, in practice we often need to deal with small to moderate sample sizes. Furthermore, if the true data generating process is unknown - as it is often the case - and the distributional characteristics of the observations may differ from period to period and subsample to subsample, a method that can deal with altering levels of heavy-tailness and skewness is needed.

In this paper, we investigate outlier detection in the case of varying distributional characteristics and in particular, for small sample sizes. Yet, as outlier detection (and treatment) involves in general a lot of subjectivity, we pursue an approach in which no (or hardly any) parameters need to be defined but that would still be able to deal with data from different non-normal distributions.

Our starting point is the adjusted boxplot proposed by Hubert and Vandervieren (2008) as it is easy to understand and compute. Also its graphical representation is an advantage. In the first step, we examine the small sample properties of this method. A well-known issue in outlier analysis is that the percentage of identified outliers is higher for small samples as shown for instance by Hoaglin et al. (1986). Unsurprisingly, we find this to be the case also with the adjusted boxplot of Hubert and Van Der Veeken (2008). Furthermore, we reveal that in small samples from right skewed distributions, the average left percentage of outliers tend to exceed the average right percentage implying that more observations are labelled as outliers in the short tail than in the long tail. This is obviously problematic. Moreover, in order to ensure that in the case of symmetric but fat-tailed data, not too many observations are declared as outliers as usually is the case, the acceptance interval needs to be adjusted for heavy-tailness.

With our modification, we address these three issues to widen the practical applications of the adjusted boxplot. Within a simulation study we show that the modified boxplot is applicable also for heavy-tailed data and small samples. We also present real data examples.

The remainder of this paper is organized as follows. In the next section, we describe the existing methods and address the issues with the adjusted boxplot in a simulation study. In section 3 we propose modifications for the adjusted boxplot and demonstrate with simulations their performance. Section 4 offers real data applications. Finally, section 5 concludes.

5.2 Existing methods

Resistant rules, firstly proposed by Tukey (1977), are very popular in outlier labelling. Also known as inner fences of the boxplot, the rule is defined as

$$q_{0.25} - k(q_{0.75} - q_{0.25}); y_{0.75} + k(q_{0.75} - q_{0.25}) \tag{5.1}$$

where $q_{0.25}, q_{0.75}$ denote the first and third sample quartiles and $k = 1.5$. $q_{0.75} - q_{0.25}$ is also known as *interquartile range.* The use of quartiles makes this method itself resistant to extreme values, unlike methods based on classical estimates of location and scale such as mean and standard deviation.

The average percentage of outside observations, that is, observations outside this interval can be derived as follows:

$$1 - (F(q_{0.75} + 1.5(q_{0.75} - q_{0.25})) - F(q_{0.25} - 1.5(q_{0.75} - q_{0.25})))$$

where F denotes the (empirical) cumulative distribution function. For normal distribution, 0.7 % of observations or 0.35 % on each side are identified as outlying as illustrated in Figure 5.1. This implies that in a large Gaussian sample only very few observations would be identified as outliers. However, in small samples this fraction is considerably higher as discussed in Hoaglin et al. (1986). This is because the quantile estimation for small samples can be problematic. Yet, obviously the parameter k can also be chosen to take greater values such as $k = 2$ or $k = 3$ in which case the fraction of outlying observations would be even smaller. However, as $k = 1.5$ is the mostly applied specification, we choose to use that in our analysis.

Although there is a considerable amount of literature on outlier detection as summarized by Barnett and Lewis (1994) for instance, approaches assessing outlying observations in the case of asymmetric or fat-tailed data are still few. One of the few works dealing with skewed data, is the work of Hubert and Vandervieren (2008) who present an adjustment to the standard boxplot to account for the potential skewness. Bruffaerts et al. (2014) propose, in turn, a generalized version of the boxplot for both skewed and heavy-tailed distributions based on the *Tukey g-and-h distribution.* The work of Xu et al. (2014) is based on a similar idea as they propose also an outlier detection approach relying on the Tukey's g-and-h distribution.

Next, we explain and discuss the adjusted boxplot of Hubert and Vandervieren (2008). The techniques based on the Tukey's g-and-h distribution we leave out as the method of Bruffaerts et al. (2014) requires multiple transformations which are hard to follow whereas the approach proposed by Xu et al. (2014) is based on the fitted distribution parameters and thus, not ideal for small samples.

5.2.1 Adjusted Boxplot

Hubert and Vandervieren (2008) introduced an *adjusted boxplot* for skewed data. The whiskers of the boxplot are adjusted to incorporate skewness as follows

$$\begin{aligned} w_l = q_{0.25} - 1.5e^{-aMC} IQR; w_u = q_{0.75} + 1.5e^{bMC} IQR \quad \text{for } MC \geq 0 \\ w_l = q_{0.25} - 1.5e^{-bMC} IQR; w_u = q_{0.75} + 1.5e^{aMC} IQR \quad \text{for } MC < 0 \end{aligned} \tag{5.2}$$

where $a = 4$, $b = 3$ and MC equals to the *medcouple*, a measure of skewness introduced by Brys et al. (2004) defined as

$$MC_n = \underset{y_i < med_n < y_j}{med} h(y_i, y_j)$$

where $q_{0.5}$ is the sample median and for all $y_i \neq y_j$ and

$$h(y_i, y_j) = \frac{(y_j - q_{0.5}) - (q_{0.5} - y_i)}{y_j - y_i}.$$

The parameters are defined such that the adjusted boxplot detect 0.7% observations as outliers as the standard boxplot does in the case of a normal distribution. The advantage of this adjustment over the standard boxplot is illustrated in Figure 5.4 for the χ^2_5- distribution. Since the adjusted boxplot does not incorporate any measure of heavy-tailness, it cannot account for heavy-tailness as noted by the authors.

We examine first the small sample properties of this method. For this purpose we simulate data from normal, χ^2- (right skewed) and t-distributions (heavy-tailed). Within each simulation, we generate 1000 samples of sizes 30, 50, and 1000 observations. We calculate the average total percentage of outside observations, as well as the average percentage of lower and upper outside observations in each case. The results are documented in Tables 5.1, 5.2 and 5.3. in the case of normal distribution, $N(0, 1)$ and $N(0, 5)$, the adjusted boxplot detects approximately as many outlying observations in a large sample ($n = 1000$) as a standard boxplot would do (Table 5.1). Also the average right and left percentage are close to each other. However, for the smaller samples of 30 or 50 observations, the average percentage increases considerably. For skewed data in Table 5.2, the adjusted boxplot works well for large samples. Yet again, the number of observations declared as outliers increases greatly for small samples. Moreover, the lower percentage becomes even higher than the upper one, which is not desirable. In Table 5.3 far too many observations are outside the fences in the case of t-distributions, this is expected since there is no adjustment for the heavy tails. These findings indicate that the adjusted boxplot of Hubert and Vandervieren (2008) tend not to perform well when applied to small samples from normal or skewed distributions or large samples from heavy-tailed distributions.

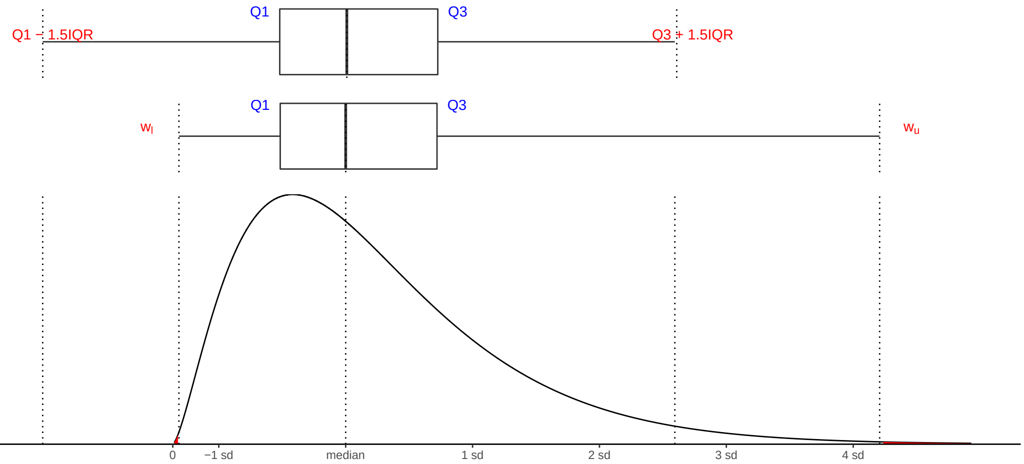

Figure 5.4: Standard and adjusted boxplot for χ^2_5-distribution

5.3 Modified adjusted boxplot

As we showed in the previous section, there are three main issues with the adjusted boxplot. First, like for standard boxplot, too many observations are declared as outliers in small samples (the total percentage exceeds 0.7% considerably). Second, for skewed data, the percentage of lower outliers exceeds the percentage of upper outliers in small samples. Third, too many observations are outside the fences in the case the tails are heavier than with a normal distribution. We aim to address these problems by modifying the adjusted boxplot and examine its performance in a simulation study.

The first issue is well-known and examined in Hoaglin et al. (1986) for standard boxplots, for instance. They show that the fraction of observations labelled as outliers increase approximately with c_r/n with $c_0 = 17$, $c_1 = 39$, $c_2 = 22$ and $c_3 = 30$, where r is the remainder from division $n/4$. Thus, we modify the fences of the adjusted boxplot by the factor c_r/n for $n < 500$.

The second point is more difficult to solve. It implies that the defined skewness adjustment factor -4 and 3 or -3 and 4, respectively, should depend on the sample size n or to be adjusted for small samples through another factor which depends on the sample size. As we found that the skewness adjustment tend to be too strong in small samples, the adjustment factor should be lower for smaller data sets. We aim to introduce a further factor as a function of the sample size $f_{ss}(n)$ to achieve better results for data sets with small number of observations. This factor should converge to one in the case of multiplicative factor or zero in the case of additive factor. We examine following multiplicative model:

$$f_{ss}(n) = 1 - n^{\alpha}$$

which should satisfy (for $mc > 0$)

$$q_{0.25} - 1.5e^{-4MC(1-n^{\alpha})} \approx q_l$$

or

$$q_{0.75} - 1.5e^{3MC(1-n^{\alpha})} \approx q_u$$

where q_p denotes the pth quantile of the sample and $l = 0.0035$ and $u = 0.9965$. Therefore, the α should satisfy

$$\ln\left(-\frac{\ln(\frac{2/3(q_{0.25}-q_l+c_r/n)}{IQR})}{-4MC}+1\right) \approx \alpha \ln(n) \tag{5.3}$$

Through a simulation study, we aim to find out the optimal value for the constant α. We generate 100 samples of sizes 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500 from various right skewed distributions (χ^2_2, χ^2_6, χ^2_{10}, $F(10, 80)$,

$F(10,90)$, $Pareto(0.5,1)$, $Pareto(1,1)$). We apply linear regression without intercept to estimate α. The coefficient estimate is found to be -0.32 (see, Table C.1 in C Appendix), which we round to -0.3 for convenience reasons. Therefore, we propose to modify the skewness adjustment by the factor $1-n^{-0.3}$ for $n<500$.

To address the third weakness, we need to adjust the fences by a tail measure. Here, we apply an index of (upper) tail weight defined as

$$TI_u = \frac{F^{-1}(0.98) - F^{-1}(0.5)}{F^{-1}(0.75) - F^{-1}(0.5)} \Big/ \frac{\Phi^{-1}(0.98) - \Phi^{-1}(0.5)}{\Phi^{-1}(0.75) - \Phi^{-1}(0.5)} \tag{5.4}$$

where F is the empirical cumulative distributive function and Φ is the distribution function of a standard normal. This means that the tail weight is relative to the normal distribution, i.e., equal to 1 in the case F is normal. As this measure quantifies the upper tail weight, we also calculate the same index for the lower tail and then take the average of these two measures as total index:

$$TI_l = \frac{F^{-1}(0.5) - F^{-1}(0.02)}{F^{-1}(0.5) - F^{-1}(0.25)} \Big/ \frac{\Phi^{-1}(0.5) - \Phi^{-1}(0.02)}{\Phi^{-1}(0.5) - \Phi^{-1}(0.25)} \tag{5.5}$$

$$TI = \frac{TI_u + TI_l}{2}. \tag{5.6}$$

To make sure that this correction works only for heavy-tailed distribution, we introduce the following condition $TI_l > 1, TI_u > 1$. This means that both tails need to be heavier than the normal case, for this correction factor to be applied. This assures that the acceptance interval is adjusted only for heavy tailed data sets and not for skewed data sets.

To find the appropriate constant for this adjustment TI^{γ}, we proceed the same way as in the previous step. We generate heavy-tailed data (100 samples with 1000 observations from $t(5)$, $t(10)$, $t(20)$, $Cauchy(0,0.5)$, $Cauchy(0,1)$, $Cauchy(0,2)$) and then, conduct the regression analysis:

$$\ln\left(2/3\frac{q_{0.25} - q_l + c_r/n}{IQR} - (-4MC(1-n^{-0.3})\right)/\ln(TI) \approx \gamma \tag{5.7}$$

The results indicate that γ equal to 2.5 (rounded) is appropriate (see, Table C.2 in C Appendix).

The final form of the modified adjusted boxplot can be written as follows

$$\begin{aligned} w_l^m &= q_{0.25} - (1.5e^{-4(1-n^{-0.3})MC} IQR \times TI^{2.5} + c_r/n); \\ w_u^m &= q_{0.75} + (1.5e^{3(1-n^{-0.3})MC} IQR \times TI^{2.5} + c_r/n) \quad \text{for } MC \geq 0 \\ w_l^m &= q_{0.25} - (1.5e^{-3(1-n^{-0.3})MC} IQR \times TI^{2.5} + c_r/n); \\ w_u^m &= q_{0.75} + (1.5e^{4(1-n^{-0.3})MC} IQR \times TI^{2.5} + c_r/n) \quad \text{for } MC < 0 \end{aligned} \tag{5.8}$$

where $TI = 1$ if $TI_l < 1$ or $TI_u < 1$. As the general small sample correction as well as the sample-size adjustment for the skewness are applied for $n < 500$, the modified boxplot performs asymptotically similar as the adjusted boxplot for large normal and skewed samples.

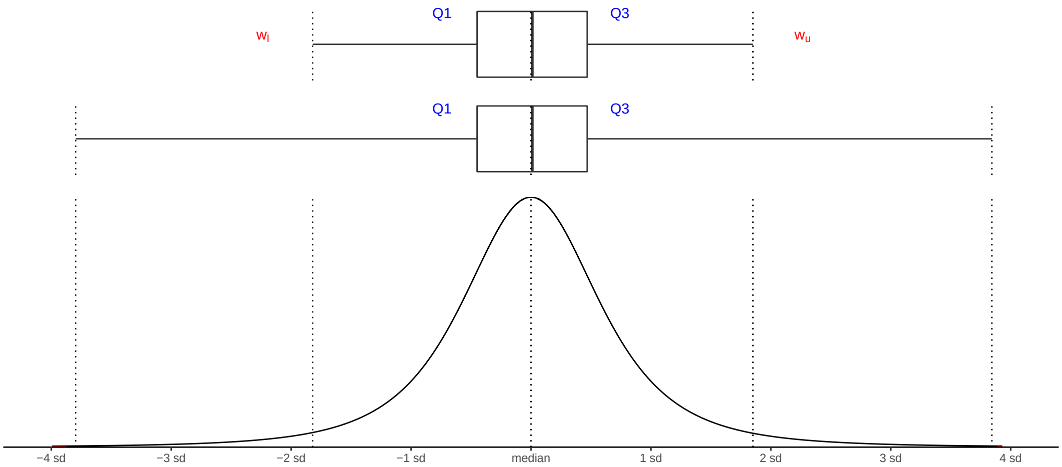

Figure 5.5: Adjusted and modified adjusted boxplot for $t(3)$-distribution

Figure 5.5 illustrates the difference between the adjusted and the modified adjusted boxplot for the $t(3)$-distribution. The whiskers of the modified version are considerably longer and therefore, less observations will be declared as outlying.

In order to examine how this modified approach behave under various models, we conduct a simulation study. We generate both uncontaminated samples like in the previous section as well as contaminated samples.

5.3.1 Uncontaminated samples

Within each simulation, we generate 1000 samples of sizes 30, 50, and 1000 observations from normal distributions ($N(0,1)$ and $N(0,5)$), chi-squared distributions (χ^2_3 and χ^2_5) and t-distributions ($t(3)$ and $t(5)$) as in the previous section.

In Table 5.1 the results for samples from normal distribution are presented. We can observe that the average percentages are very similar for both methods for large samples. Yet, for small samples, the average share of outliers is for the modified boxplot considerably lower, due to the small sample correction factor. This is also illustrated in Figure 5.6.

For right skewed data (Table 5.2) the asymptotic properties of these both approaches are approximately the same. However, again, differences occur for smaller samples. As in the previous case, the total percentage is higher for the adjusted boxplot than for the modified version. Moreover, whereas for the standard adjusted boxplot the lower percentage becomes higher than the upper percentage, this is not the case for the modified boxplot. This implies that the introduced skewness adjustment for small samples does its intended job. Figure 5.7 demonstrates this case graphically.

As Hubert and Vandervieren (2008) do not incorporate any heavy-tailness measure in their adjusted boxplot, it labels too many observations as outlying in samples from t-distribution as shown in Table 5.3. Yet, the modified version can deal with the heavy tails better as it identifies about the same number of outliers than in the case of a normal distribution.

Table 5.1: Average percentage of outside observations for normal distribution

		Standard Boxplot			Adjusted Boxplot			Modified Boxplot		
n	SD	Tot %	Low %	Up %	Tot %	Low %	Up %	Tot %	Low %	Up %
30	1	1.69	0.85	0.84	4.56	2.20	2.36	0.35	0.14	0.21
50	1	1.15	0.56	0.59	3.32	1.59	1.73	0.60	0.28	0.32
1000	1	0.71	0.35	0.35	0.90	0.46	0.44	0.80	0.41	0.39
30	5	1.41	0.71	0.70	4.30	2.35	1.95	1.43	0.78	0.65
50	5	1.26	0.64	0.62	3.43	1.60	1.83	1.38	0.68	0.70
1000	5	0.71	0.36	0.36	0.93	0.47	0.46	0.81	0.41	0.40

Table 5.2: Average percentage of outside observations for χ^2 distribution

		Standard Boxplot			Adjusted Boxplot			Modified Boxplot		
n	DF	Tot %	Low %	Up %	Tot %	Low %	Up %	Tot %	Low %	Up %
30	3	4.07	0.00	4.07	4.15	2.75	1.41	1.24	0.00	1.24
50	3	3.89	0.00	3.89	3.18	2.06	1.12	1.23	0.02	1.21
1000	3	3.77	0.00	3.77	0.42	0.03	0.38	0.42	0.03	0.38
30	5	3.03	0.00	3.02	4.16	2.59	1.57	1.28	0.07	1.22
50	5	2.95	0.00	2.95	3.28	2.06	1.21	1.24	0.08	1.16
1000	5	2.79	0.00	2.79	0.62	0.18	0.44	0.62	0.18	0.44

Note: DF = Degrees of freedom

Table 5.3: Average percentage of outside observations for t-distribution

		Standard Boxplot			Adjusted Boxplot			Modified Boxplot		
n	DF	Tot %	Low %	Up %	Tot %	Low %	Up %	Tot %	Low %	Up %
30	3	5.82	2.93	2.89	7.77	3.88	3.88	1.22	0.64	0.58
50	3	5.72	2.94	2.78	7.31	3.78	3.53	1.73	0.90	0.83
1000	3	5.53	2.76	2.76	5.73	2.87	2.86	0.67	0.33	0.34
30	5	3.80	1.87	1.93	6.25	3.10	3.14	0.77	0.36	0.41
50	5	3.36	1.71	1.64	5.20	2.78	2.42	1.39	0.72	0.67
1000	5	3.32	1.65	1.66	3.54	1.75	1.79	0.79	0.40	0.39

Note: DF = Degrees of freedom

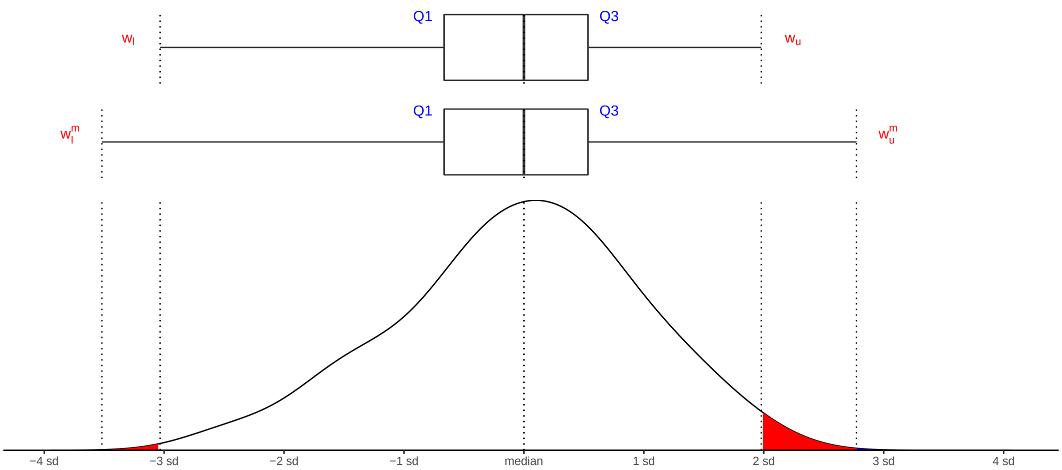

Figure 5.6: Adjusted and modified adjusted boxplot for a small normal sample (n = 30)

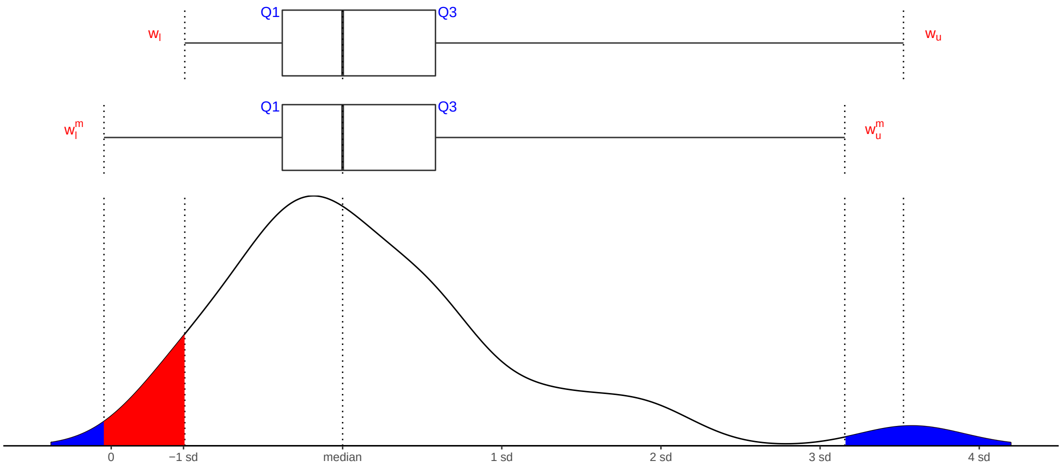

Figure 5.7: Adjusted and modified adjusted boxplot for a small sample (n = 30) from χ^2_5-distribution

5.3.2 Contaminated samples

To examine how robust the proposed version of the adjusted boxplot is, we also generate data samples from *contaminated distributions*, firstly introduced by **?** and since then widely applied in outlier and robust analysis to simulate samples with potentially outlying observations. The samples are drawn from contaminated distributions defined as

$$F = (1 - \lambda)G + \lambda H \tag{5.9}$$

where G is the original distribution, λ denotes the contamination level and H is the distribution from which the outliers are drawn. Since we are concerned about outliers in the case in which the true distribution itself is asymmetric or heavy-tailed, we generate data first from

$$F = (1 - \lambda)\chi^2_5 + \lambda N(30, 1)$$

and let the λ vary between 0 and 0.05. This corresponds to right skewed data with asymmetric contamination from a normal sample.

The second group of samples are drawn from

$$F = (1 - \lambda)t_7 + \lambda t_1$$

where the contamination level λ is again let to vary from 0 to 0.05, i.e., we contaminate a heavy tailed distribution with outliers from a distribution with even heavier tails.

For samples from an asymmetrically contaminated χ^2-distribution we observe that the modified method do detect more outliers for higher levels of contamination like the adjusted boxplot does. Particularly, more upper outliers are identified than lower ones, as expected since we contaminated the sample asymmetrically with only upper contaminants. Moreover, the lower percentage is always smaller than the upper percentage for the modified boxplot which indicates that the incorporated skewness correction factor for small samples performs as requested.

Table 5.4: Average percentage of outside observations for contaminated χ^2-distribution

		Standard Boxplot			Adjusted Boxplot			Modified Boxplot		
n	CL	Tot %	Low %	Up %	Tot %	Low %	Up %	Tot %	Low %	Up %
30	0.000	3.12	0.00	3.11	4.36	2.67	1.69	1.29	0.05	1.24
50	0.000	3.00	0.00	3.00	2.95	1.84	1.11	1.18	0.09	1.09
1000	0.000	2.81	0.00	2.81	0.60	0.16	0.44	0.60	0.16	0.44
30	0.025	4.77	0.01	4.77	5.92	2.89	3.02	2.96	0.01	2.94
50	0.025	4.85	0.00	4.85	5.20	2.26	2.95	3.20	0.09	3.11
1000	0.025	4.72	0.00	4.72	3.06	0.34	2.72	3.06	0.34	2.72
30	0.050	6.74	0.00	6.74	7.36	3.28	4.07	4.55	0.02	4.53
50	0.050	6.72	0.00	6.72	6.86	2.56	4.30	5.11	0.09	5.02
1000	0.050	6.68	0.00	6.68	5.68	0.58	5.09	5.68	0.58	5.09

Note: CL = Contamination level

in the case of a contaminated t-distribution (Table 5.5), the modified boxplot labels more outliers the higher the contamination level, like the adjusted boxplot. Yet, the average level outlying observations is considerably lower due to adjustment for heavy tails.

All in all, the modified adjusted boxplot is found to perform well also in the contaminated samples, i.e., it is able to identify an appropriate share of outlying observations but does not declare too many observations as outliers.

Table 5.5: Average percentage of outside observations for contaminated t-distribution

		Standard Boxplot			Adjusted Boxplot			Modified Boxplot		
n	CL	Tot %	Low %	Up %	Tot %	Low %	Up %	Tot %	Low %	Up %
30	0.000	3.15	1.59	1.56	5.71	2.84	2.87	0.72	0.38	0.34
50	0.000	2.74	1.30	1.44	4.67	2.17	2.50	1.15	0.52	0.63
1000	0.000	2.52	1.26	1.26	2.75	1.37	1.39	0.83	0.42	0.41
30	0.025	3.58	1.86	1.72	6.12	2.98	3.14	0.84	0.49	0.34
50	0.025	3.14	1.59	1.55	5.24	2.65	2.59	1.42	0.72	0.70
1000	0.025	2.95	1.48	1.47	3.16	1.58	1.57	0.97	0.48	0.49
30	0.050	3.70	1.79	1.91	6.20	3.22	2.98	0.98	0.48	0.50
50	0.050	3.63	1.82	1.81	5.68	2.92	2.76	1.55	0.78	0.77
1000	0.050	3.32	1.67	1.66	3.54	1.79	1.75	1.13	0.57	0.56

Note: CL = Contamination level

5.4 Real-data examples

We illustrate the usefulness of the modified adjusted boxplot within two real data applications.

5.4.1 Heavy-tailed data

In the first example, we use financial market data as they traditionally display heavy tails. We use the daily values of Dow Jones Industrial Average (DJIA) from 15th October 2007 to 30th September 2017 and calculate the daily returns. In Figure 5.8 we show the kernel density of the data and the adjusted as well as the modified adjusted boxplots. It is obvious that the data has relatively heavy tails. That is why also the whiskers of the modified boxplot are considerably longer than those of the adjusted boxplot of Hubert and Vandervieren (2008) and thus, labelling much fewer observations as outlying than the adjusted boxplot (1.95% vs. 9.04 %).

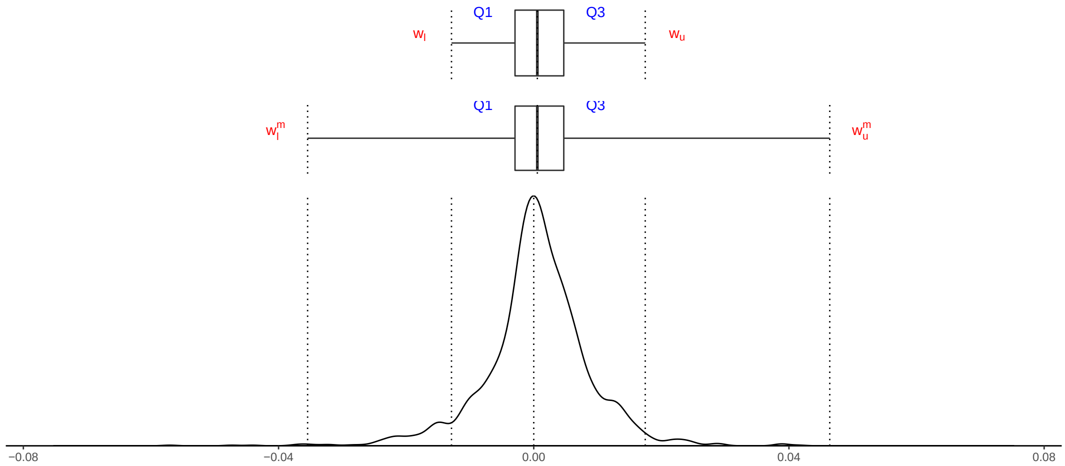

Figure 5.8: Adjusted and modified adjusted boxplot for Dow Jones Industrial Average

5.4.2 Small skewed sample

The second example covers the case of a small skewed sample. For this application, we consider the coal mining disaster data of Jarrett (1979). We calculate the time differences between the disaster (in days) and use the first 30 observations only. In Figure 5.9 the differences between the boxplot methods are illustrated. The lower whisker of the modified version is longer and thus, ensuring that fewer observations in the short tail seem to be outlying. On the other hand, the left whisker is shorter than that of the adjusted boxplot and therefore, the upper percentage of outlying observations turns also to be higher.

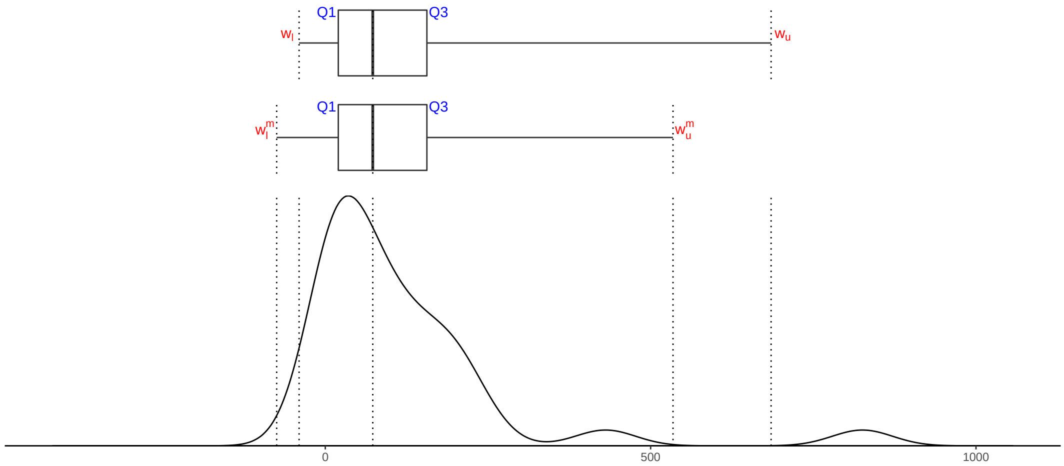

Figure 5.9: Adjusted and modified adjusted boxplot for a small sample (n = 30) from coal data

5.5 Conclusions

Even though a large number of techniques to identify outlying observations have already been proposed, methods to deal with outliers in skewed or heavy-tailed data are still scarce. Hubert and Vandervieren (2008) propose an adjusted boxplot for skewed distributions as the standard boxplot often incorrectly labels too many observations as outliers. However, we show that their method is not optimal for small normal or skewed samples nor for heavy tailed data. First, too many observations are declared as outliers in small samples. Second, for skewed data, the percentage of lower outliers exceeds the percentage of upper outliers in small samples. Third, too many observations are outside the fences in the case the tails are heavier than in the case of a normal distribution.

In this paper, we propose a modification of the adjusted boxplot of Hubert and Vandervieren (2008) to address these three issues. With our modification, we widen the practical application of the adjusted boxplot method. Within a simulation study and real data examples we show that the modified adjusted boxplot is applicable also for heavy-tailed data and for small samples from skewed distributions.

Bibliography

Aielli, G. P. (2013). Dynamic Conditional Correlation: on Properties and Estimation. *Journal of Business & Economic Statistics*, 31(3):282–299.

Antonakakis, N., Chatziantoniou, I., and Filis, G. (2013). Dynamic co-movements of stock market returns, implied volatility and policy uncertainty. *Economics Letters*, 120:87–92.

Barnett, V. and Lewis, T. (1994). *Outliers in statistical data.* Wiley, New York, 3. edition.

Beaudry, P. and Lucke, B. (2010). Letting Different Views about Business Cycles Compete. *NBER Macroeconomic Annual*, 24(1):413–456.

Beaudry, P., Nam, D., and Wang, J. (2011). Do mood swings drive business cycles and is it rational? Working paper no. 17651, National Bureau of Economic Research.

Beaudry, P. and Portier, F. (2004). An exploration into Pigou's theory of cycles. *Journal of Monetary Economics*, 51(6):1183–1216.

Beaudry, P. and Portier, F. (2006). Stock prices, news, and economic fluctuations. *American Economic Review*, 96(4):1293–1307.

Beaudry, P. and Portier, F. (2014). News-Driven Business Cycles: Insights and Challenges. *Journal of Economic Literature*, 52(4):993–1074.

Beveridge, W. H. (1909). *Unemployment : a problem of industry.* Longmans Green, London.

Billio, M., Caporin, M., and Gobbo, M. (2006). Flexible Dynamic Conditional Correlation multivariate GARCH models for asset allocation. *Applied Financial Economics Letters*, 2(2):123–130.

Boldin, M. and Wright, J. H. (2015). Weather-Adjusting Economic Data. *Brookings Papers on Economic Activity*, 2:227–278.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics*, 72:498–505.

Bollerslev, T., Engle, R. F., and Wooldridge, J. M. (1988). A Capital Asset Pricing Model with Time-Varying Covariances. *Journal of Political Economy*, 96(1):116–131.

Bruffaerts, C., Verardi, V., and Vermandele, C. (2014). A generalized boxplot for skewed and heavy-tailed distributions. *Statistics and Probability Letters*, 95:110–117.

Brys, G., Hubert, M., and Struyf, A. (2004). A robust measure of skewness. *Journal of Computational and Graphical Statistics*, 13(4):996–1017.

Burns, A. F. and Mitchell, W. C. (1946). *Measuring business cycles.*

Busse, M. R., Pope, D. G., Pope, J. C., and Silva-Risso, J. (2015). The psychological effect of weather on car purchases. *Quarterly Journal of Economics*, 130(1):371–414.

Cappiello, L., Engle, R. F., and Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics*, 4(4):537–572.

Carvalho, V. M. (2008). *Aggregate fluctuations and the network structure of intersectoral trade.* PhD thesis, The University of Chicago.

Cassou, S. P. and Vázquez, J. (2014). Employment comovements at the sectoral level over the business cycle. *Empirical Economics*, 46(4):1301–1323.

Cermeno, R. and Grier, K. B. (2008). Conditional Heteroskedasticity and Cross-Sectional Dependence in Panel Data: An Empirical Study of Inflation Uncertainty in the G7 countries. In B.H., B., editor, *Panel Data Econometrics*, chapter 10, pages 259–277.

Chambers, R. L. (1986). Outlier robust finite population estimation. *Journal of the American Statistical Association*, 81(396):1063–1069.

Christiano, L. J. and Fitzgerald, T. J. (1998). The business cycle: It's still a puzzle. *Economic Perspectives*, 22:56–83.

Clark, J. M. (1934). *Strategic Factors in Business Cycles.* National Bureau of Economic Research.

Cooper, R. and Haltiwanger, J. (1990). Inventories and the propagation of sectoral shocks. *The American Economic Review*, 80(1):170–190.

Ding, Z., Granger, C. W., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1):83–106.

Dupor, B. (1999). Aggregation and irrelevance in multi-sector models. *Journal of Monetary Economics*, 43(2):391–409.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4):987–1007.

Engle, R. F. (2002). Dynamic Conditional Correlation - a simple class of multivariate GARCH models. *Journal of Business & Economic Statistics*, 20(3):339–350.

Engle, R. F. (2009). *Anticipating correlations: a new paradigm for risk management.* Princeton University Press.

Engle, R. F. and Kelly, B. T. (2012). Dynamic Equicorrelation. *Journal of Business & Economic Statistics*, 30(2):212–228.

Engle, R. F. and Kroner, K. F. . (1995). Multivariate Simultaneous Generalized Arch. *Econometric Theory*, 11(1):122–150.

Engle, R. F., Shephard, N., and Sheppard, K. (2007). Fitting and testing vast dimensional time-varying covariance models. Working paper no. fin-07-046, New York University.

Engle, R. F. and Sheppard, K. (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. Working paper no. 8554, National Bureau of Economic Research.

Fama, E. F. (1990). Stock returns, expected returns, and real activity. *The Journal of Finance*, 45(4):1089–1108.

Foerster, A. T., Sarte, P.-D. G., and Watson, M. W. (2011). Sectoral versus Aggregate Shocks: A Structural Factor Analysis of Industrial Production. *The Journal of Political Economy*, 119(1):1–38.

Forni, M., Gambetti, L., and Sala, L. (2014). No News in Business Cycles. *The Economic Journal*, 124(581):1168–1191.

Fuhrer, J. C. (1993). What role does consumer sentiment play in the U.S. macroeconomy? *New England Economic Review*, pages 32–44.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, 48(5):1779–1801.

Ho, K.-Y., Tsui, A. K., and Zhang, Z. (2009a). Volatility dynamics of the UK business cycle: A multivariate asymmetric garch approach. *Economie internationale*, 117(1):31–46.

Ho, K. Y., Tsui, A. K., and Zhang, Z. (2009b). Volatility dynamics of the US business cycle: A multivariate asymmetric GARCH approach. *Mathematics and Computers in Simulation*, 79(9):2856–2868.

Hoaglin, D. C., Iglewicz, B., and Tukey, J. W. (1986). Performance of Some Resistant Rules for Labeling Outlier. *Journal of the American Statistical Association*, 81(396):991–999.

Hornstein, A. (2000). The business cycle and industry comovement. *Economic Quarterly-Federal Reserve Bank of Richmond*, 86(1):27–48.

Horvath, M. (2000). Sectoral shocks and aggregate fluctuations. *Journal of Monetary Economics*, 45(1):69–106.

Horvath, M. T. K. (1998). Cyclicality and Sectoral Linkages: Aggregate Fluctuations from Independent Sectoral Shocks. *Review of Economic Dynamics*, 1(4):781–808.

Hubert, M. and Van Der Veeken, S. (2008). Outlier detection for skewed data. *Journal of Chemometrics*, 22(3-4):235–246.

Hubert, M. and Vandervieren, E. (2008). An adjusted boxplot for skewed distributions. *Computational Statistics and Data Analysis*, 52(12):5186–5201.

Inklaar, R., Jong-A-Pin, R., and de Haan, J. (2008). Trade and business cycle synchronization in OECD countries-A re-examination. *European Economic Review*, 52(4):646–666.

Jaimovich, N. and Rebelo, S. (2009). Can news about the future drive the business cycle? *American Economic Review*, 99(4):1097–1118.

Jarrett, R. G. (1979). A Note on the Intervals Between Coal-Mining Disasters. *Biometrika*, 66(1):191–193.

Jondeau, E. and Rockinger, M. (2006). The Copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance*, 25(5):827–853.

Kawasaki, S. and Zimmermann, K. F. (1986). Testing the Rationality of Price Expectations for Manufacturing Firms. *Applied Economics*, 18(12):1335–1347.

Lamla, M. J., Lein, S. M., and Sturm, J.-E. (2007). News and sectoral comovement. Working paper no. 183, KOF Swiss Economic Institute.

Lee, J. (2006). The comovement between output and prices: Evidence from a dynamic conditional correlation GARCH model. *Economics Letters*, 91(1):110–116.

Linden, F. (1962). Consumer markets: Merchandising Weather. *The Conference Board Business Record*, 19(6):15–16.

Long, J. B. and Plosser, C. I. (1987). Sectoral vs. Aggregate Shocks In The Business Cycle. *The American Economic Review*, 77(2):333–336.

Long, J. B. J. and Plosser, C. I. (1983). Real Business Cycles. *Journal of Political Economy*, 91(1):39–69.

Lucas, R. E. (1977). Understanding business cycles. In *Carnegie-Rochester conference series on public policy*, volume 5, pages 7–29. North-Holland.

Mankiw, N. G. (1985). Consumer durables and the real interest rate. *The Review of Economics and Statistics*, 67(3):353–362.

Maunder, W. (1973). Weekly weather and economic activities on a national scale: an example using United States retail trade data. *Weather*, 28(1):2–19.

Murray, K. B., Di Muro, F., Finn, A., and Popkowski Leszczyc, P. (2010). The effect of weather on consumer spending. *Journal of Retailing and Consumer Services*, 17(6):512–520.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59(2):347–370.

Papageorgiou, T., Michaelides, P. G., and Milios, J. G. (2010). Business cycles synchronization and clustering in Europe (1960-2009). *Journal of Economics and Business*, 62(5):419–470.

Pigou, A. C. (1927). *Industrial Fluctuations*. Macmillan, London.

Raddatz, C. (2010). Credit Chains and Sectoral Comovement: Does the Use of Trade Credit Amplify Sectoral Shocks? *The Review of Economics and Statistics*, 92(4):985–1003.

Rebelo, S. (2005). Real business cycle models: Past, present and future. *Scandinavian Journal of Economics*, 107(2):217–238.

Romer, C. D. (1991). The cyclical behavior of individual production series, 1889-1984. *The Quarterly Journal of Economics*, 106(1):1–31.

Sandqvist, A. P. (2017). Dynamics of sectoral business cycle comovement. *Applied Economics*, 49(47):4742–4759.

Schmitt-Grohe, S. and Uribe, M. (2012). What's News in Business Cycles. *Econometrica*, 80(6):2733–2764.

Shea, J. (2002). Complementarities and comovements. *Journal of Money, Credit, and Banking*, 34(2):412–433.

Starr-McCluer, M. (2000). *The effects of weather on retail sales.* Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board.

Tse, Y. K. and Tsui, A. K. C. (2002). A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model With Time-Varying Correlations. *Journal of Business & Economic Statistics*, 20(3):351–362.

Tukey, J. W. (1977). *Exploratory Data Analysis.* Addison-Wesley, Reading, MA.

Veldkamp, L. and Wolfers, J. (2007). Aggregate shocks or aggregate information? Costly information and business cycle comovement. *Journal of Monetary Economics*, 54:37–55.

Xu, Y., Iglewicz, B., and Chervoneva, I. (2014). Robust estimation of the parameters of g-and-h distributions, with applications to outlier detection. *Computational Statistics & Data Analysis*, 75:66–80.

Appendix A

Appendix Chapter 2

Table A.1: US data definition

NAICS_Code	Description	Short Label
311	Nondurable Goods: Food	Food
312	Nondurable Goods: Beverage and tobacco product	BevTob
313	Nondurable Goods: Textile mills	Textile
314	Nondurable Goods: Textile product mills	Textileprod
315	Nondurable Goods: Apparel	Apparel
316	Nondurable Goods: Leather and allied product	Leather
321	Durable manufacturing: Wood product	Wood
322	Nondurable manufacturing: Paper	Paper
323	Nondurable manufacturing: Printing and related support activities	Print
324	Nondurable manufacturing: Petroleum and coal products	Petroleum
325	Nondurable manufacturing: Chemical	Chemical
326	Nondurable manufacturing: Plastics and rubber products	Rubber
327	Durable manufacturing: Nonmetallic mineral product	Mineral
331	Durable manufacturing: Primary metal	MetalPri
332	Durable manufacturing: Fabricated metal product	MetalFab
333	Durable manufacturing: Machinery	Machinery
334	Durable manufacturing: Computer and electronic product	Electronics
335	Durable manufacturing: Electrical equipment, appliance, and component	Electricals
336	Durable Goods: Transportation equipment	Transportation
337	Durable manufacturing: Furniture and related product	Furniture
339	Durable manufacturing: Miscellaneous	Miscellaneous

Table A.3: Diagnostic test for ARMA(l,k)-residuals for US data

	w LB (20)	w LB^2 (20)	Arch LM (10)	Skewness	Kurtosis	Jarque-Bera
Food	21.12	101.49 ***	27.08 ***	-0.06	3.91	18.12 ***
BevTob	26.31 **	74.37 ***	33.16 ***	-0.04	3.69	10.43 ***
Textile	34.27 ***	229.95 ***	89.09 ***	-0.01	5.55	140.41 ***
Textileprod	15.54	33.68 ***	22.56 **	0	4.05	23.89 ***
Apparel	26.34 *	86.67 ***	65.32 ***	-0.05	4.88	76.38 ***
Leather	14.24	37.59 ***	9.72	-0.62	7.22	418.74 ***
Wood	23.16 *	16.79	5.66	-1.68	18.10	5171.38 ***
Paper	40.53 ***	154.64 ***	72.47 ***	-0.03	5.07	93.02 ***
Print	34.4 ***	132.69 ***	57.55 ***	-0.17	3.63	11.07 ***
Petroleum	16.17	79.53 ***	76.85 ***	-0.03	9.26	848.07 ***
Chemical	18.08	135.63 ***	92.07 ***	-0.73	8.94	810.2 ***
Rubber	41.59 ***	154.2 ***	74.02 ***	0.19	15.00	3133.99 ***
Mineral	20.42	131.21 ***	49.81 ***	-0.07	3.94	19.39 ***
MetalPri	26.72 *	78.11 ***	59.88 ***	-0.25	4.40	47.53 ***
MetalFab	29.91 **	80.24 ***	54.48 ***	-0.24	4.87	80.93 ***
Machinery	36.16 ***	33.57 ***	25.72 ***	0.01	3.62	8.27 **
Electronics	23.55	170.23 ***	68.58 ***	-0.02	5.27	111.89 ***
Electricals	50.89 ***	100.82 ***	35.85 ***	0.03	4.79	69.22 ***
Transportation	9.83	40.68 ***	34.69 ***	-0.35	9.96	1059.52 ***
Furniture	28.46 *	48.55 ***	29.66 ***	-0.05	5.41	126.27 ***
Miscellaneous	19.05	65.72 ***	35.46 ***	-0.16	4.21	33.85 ***

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A.4: Diagnostic test for ARMA(l,k)-residuals for German data

	w LB (20)	w LB^2 (20)	Arch LM (10)	Skewness	Kurtosis	Jarque-Bera
FoodBevTob	42.85 ***	25.57	8.43	0.05	2.96	0.14
Textiles	20.25	40.45 ***	28.9 ***	-0.37	5.92	109.57 ***
Apparel	11.99	18.41	7.6	-0.31	3.55	8.29 **
Leather	22.42	39.88 ***	29.75 ***	-0.23	5.47	76.81 ***
Wood	30.03 *	47.34 ***	36.96 ***	0.08	4.39	23.7 ***
Paper	31.16 **	32.14 **	17.29 *	-0.56	4.47	41.04 ***
Print	13.94	14.9	4.19	-0.08	4.16	16.66 ***
Petroleum	22.08	24.87	7.76	-0.67	4.78	60.06 ***
ChemicalsPharma	26.77	32.81 **	28.22 ***	-0.38	6.32	140.49 ***
Rubber	21	38.36 ***	25.73 ***	-1.20	11.90	1026.61 ***
GlassStone	26.39	62.98 ***	16.86 *	-0.81	7.97	330.79 ***
MetalPri	14.53	78.34 ***	61.43 ***	-0.37	4.63	38.79 ***
MetalFab	21.8	55.4 ***	47.65 ***	-0.62	5.70	106.77 ***
Electronics	12.34	7.38	5.57	-0.90	12.30	1083.45 ***
Electricals	40 ***	39.69 ***	31.6 ***	0.15	5.24	61.91 ***
Machinery	30.26 **	47.7 ***	41.45 ***	-0.36	10.70	732.1 ***
Vehicles	11.82	67.67 ***	36.9 ***	-0.35	6.16	126.92 ***
Transport	21.19	73.87 ***	32.91 ***	-0.41	3.87	17.51 ***
Furniture	31.46 **	32.23 **	31.53 ***	-0.54	8.35	360.48 ***
Other	17.63	37.2 ***	19.37 **	-0.39	3.88	16.53 ***
Repair	28.23 *	18.82	9.15	1.43	9.02	538.6 ***

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table A.2: German data definition

NACE Code	Description	Short Label
C10-C12	Manufacture of food products; beverages and tobacco products	FoodBevTob
C13	Manufacture of textiles	Textiles
C14	Manufacture of wearing apparel	Apparel
C15	Manufacture of leather and related products	Leather
C16	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	Wood
C17	Manufacture of paper and paper products	Paper
C18	Printing and reproduction of recorded media	Print
C19	Manufacture of coke and refined petroleum products	Petroleum
C20-C21	Manufacture of chemicals and chemical products; basic pharmaceutical products and pharmaceutical preparations	ChemicalsPharma
C22	Manufacture of rubber and plastic products	Rubber
C23	Manufacture of other non-metallic mineral products	GlassStone
C24	Manufacture of basic metals	MetalPri
C25	Manufacture of fabricated metal products, except machinery and equipment	MetalFab
C26	Manufacture of computer, electronic and optical products	Electronics
C27	Manufacture of electrical equipment	Electricals
C28	Manufacture of machinery and equipment n.e.c.	Machinery
C29	Manufacture of motor vehicles, trailers and semi-trailers	Vehicles
C30	Manufacture of other transport equipment	Transport
C31	Manufacture of furniture	Furniture
C32	Other manufacturing	Other
C33	Repair and installation of machinery and equipment	Repair

Table A.5: Individual ARMA-GARCH models for US data

	Food	BevTob	Textile	Textileprod	Apparel	Leather	Wood	Paper	Print	Petroleum	Chemical
mu	0.00^{***}	0.00	-0.00^{***}	0.00	-0.00^{**}	-0.00^{***}	0.00	0.00	0.00	0.00	0.00^{***}
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ar1		-0.64^{***}	-1.65^{***}	-1.19^{***}	0.81^{***}	0.04	-1.07^{***}	-0.27^{***}	0.60^{***}		0.80^{***}
		(0.04)	(0.00)	(0.09)	(0.05)	(0.05)	(0.00)	(0.06)	(0.13)		(0.06)
ar2		-0.96^{***}	-0.53^{***}	-0.97^{***}		0.14^{***}	-0.89^{***}	0.11^{**}	0.27^{***}		
		(0.02)	(0.00)	(0.05)		(0.05)	(0.00)	(0.05)	(0.05)		
ar3			0.47^{***}	0.14		0.19^{***}	0.09^{***}	0.24^{***}			
			(0.00)	(0.11)		(0.04)	(0.00)	(0.04)			
ar4			0.17^{***}	0.19							
			(0.00)	(0.24)							
ar5				0.11							
				(0.15)							
ma1	-0.18^{***}	0.18^{***}	1.83^{***}	1.15^{***}	-0.85^{***}		1.15^{***}		-0.77^{***}	-0.13^{**}	-0.71^{***}
	(0.06)	(0.07)	(0.00)	(0.11)	(0.07)		(0.00)		(0.07)	(0.05)	(0.08)
ma2		0.73^{***}	1.03^{***}	0.93^{***}	0.22^{***}		0.98^{***}				
		(0.03)	(0.00)	(0.05)	(0.07)		(0.00)				
ma3		-0.36^{***}									
		(0.05)									
omega	0.00	0.00^{***}	0.00^{**}	-2.01	0.00^{**}	-0.09^{***}	0.00^{***}	0.00	0.00	0.00	0.00^{***}
	(0.00)	(0.00)	(0.00)	(1.94)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
alpha1	0.02	0.05^{***}	0.40^{*}	-0.01	0.56^{***}	-0.10^{***}	0.05^{***}	0.01	0.08	0.01	0.19^{***}
	(0.06)	(0.01)	(0.21)	(0.15)	(0.14)	(0.03)	(0.01)	(0.12)	(0.29)	(0.04)	(0.07)
beta1	0.97^{***}	0.93^{***}	0.52^{***}	0.74^{***}	0.16	0.99^{***}	0.90^{***}	0.80^{***}	0.91^{***}	0.70^{***}	0.36^{***}
	(0.06)	(0.01)	(0.18)	(0.25)	(0.23)	(0.00)	(0.02)	(0.05)	(0.30)	(0.14)	(0.09)
gamma1				0.41^{***}		0.06^{***}		0.21^{*}		0.29^{***}	0.25
				(0.13)		(0.02)		(0.11)		(0.09)	(0.18)
skew						0.88^{***}	0.88^{***}	0.87^{***}			
						(0.06)	(0.05)	(0.06)			
shape			6.17^{***}	8.03	6.29^{***}	7.77^{***}	4.26^{***}	10.64		6.89^{***}	8.37^{***}
			(2.39)	(6.00)	(1.55)	(2.74)	(0.72)	(7.98)		(2.25)	(2.15)
Num. obs.	519.00	519.00	519.00	519.00	519.00	519.00	519.00	519.00	519.00	519.00	519.00

	Rubber	Mineral	MetalPri	MetalFab	Machinery	Electronics	Electricals	Transportation	Furniture	Miscellaneous
mu	0.00^{***}	0.00	0.00	0.00^{**}	0.00^{***}	0.01^{***}	0.00	0.00	0.00	0.00^{***}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ar1		0.61^{**}	0.03	0.17^{***}	1.83^{***}	0.93^{***}	0.85^{***}	-0.06	0.17^{***}	-0.04
		(0.28)	(0.06)	(0.04)	(0.00)	(0.02)	(0.05)	(0.07)	(0.06)	(0.04)
ar2			0.19^{***}	0.17^{***}	-0.84^{***}				0.18^{***}	0.15^{***}
			(0.04)	(0.03)	(0.00)				(0.06)	(0.04)
ar3				0.29^{***}						0.23^{***}
				(0.03)						(0.04)
ar4										
ar5										
ma1	0.14^{**}	-0.68^{**}			-1.96^{***}	-0.66^{***}	-0.93^{***}			
	(0.06)	(0.26)			(0.00)	(0.07)	(0.08)			
ma2	0.09^{*}	0.15^{***}			1.24^{***}	0.03	0.25^{***}			
	(0.05)	(0.06)			(0.00)	(0.07)	(0.04)			
ma3	0.22^{***}				-0.28^{***}					
	(0.04)				(0.00)					
omega	0.00^{***}	0.00^{***}	0.00^{***}	-0.78^{***}	0.00^{***}	-1.56	0.00	-1.36	-1.34	0.00^{***}
	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)	(1.17)	(0.00)	(1.13)	(1.11)	(0.00)
alpha1	0.04	0.06	0.45^{***}	-0.16^{***}	0.06^{***}	-0.18^{**}	0.00	-0.26^{***}	-0.18^{***}	0.32^{***}
	(0.03)	(0.09)	(0.10)	(0.01)	(0.01)	(0.07)	(0.03)	(0.06)	(0.05)	(0.08)
beta1	0.78^{***}	0.49^{***}		0.92^{***}	0.45^{***}	0.83^{***}	0.92^{***}	0.82^{***}	0.85^{***}	0.18
	(0.03)	(0.11)		(0.00)	(0.12)	(0.13)	(0.05)	(0.15)	(0.13)	(0.22)
gamma1	0.23^{***}	0.40^{***}		0.12^{***}	0.38^{***}	0.41^{***}	0.13^{**}	0.33	0.21^{*}	
	(0.07)	(0.09)		(0.04)	(0.11)	(0.12)	(0.06)	(0.23)	(0.12)	
skew	0.88^{***}	0.92^{***}		0.88^{***}				0.80^{***}		0.92^{***}
	(0.05)	(0.06)		(0.06)				(0.05)		(0.05)
shape	8.66^{**}	7.80^{***}	7.52^{***}			7.70^{***}		5.63^{***}	9.75^{**}	5.85^{***}
	(3.46)	(2.91)	(2.35)			(1.73)		(1.47)	(3.80)	(1.33)
Num. obs.	519.00	519.00	519.00	519.00	519.00	519.00	519.00	519.00	519.00	519.00

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table A.6: Univariate ARMA-GARCH models for German data

	FoodBevTob	Textiles	Apparel	Leather	Wood	Paper	Print	Petroleum	ChemicalsPharma	Rubber	GlassStone	MetalPri	MetalFab	Electronics	Electricals	Machinery	Vehicles	Transport	Furniture	Other	Repair
mu	0.00*	−0.00***	−0.01***	−0.00***	0.00	0.00*	−0.00	0.00	0.00*	0.00*	−0.00*	0.00	0.00***	0.00**	0.00*	0.00	0.00**	0.00	−0.00	0.00***	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ar1			0.81***	0.73		−0.36***	−0.55***		−0.33***	−0.30***		−0.50	−0.31***	−0.25***	−0.24***	0.83***	−0.46***	0.81***	−0.50***	−0.65***	−0.42***
			(0.04)	(0.81)		(0.05)	(0.06)		(0.06)	(0.06)		(0.83)	(0.08)	(0.06)	(0.08)	(0.06)	(0.05)	(0.11)	(0.07)	(0.06)	(0.05)
ar2			−0.87***				−0.25***						−0.09	0.11**	0.10				−0.16***	−0.31***	−0.23***
			(0.02)				(0.05)						(0.07)	(0.05)	(0.08)				(0.05)	(0.05)	(0.05)
ar3													0.12***	0.29***	0.22***						
													(0.04)	(0.04)	(0.05)						
ar4													0.15***								
													(0.06)								
ma1	−0.66***	−0.48***	−1.30***	−1.29	−0.35***			−0.46***			−0.21***	0.18				−1.36***		−1.31***			
	(0.05)	(0.08)	(0.03)	(0.79)	(0.07)			(0.08)			(0.06)	(0.80)				(0.07)		(0.12)			
ma2		0.06	1.15***	0.46				−0.23***				−0.07				0.56***		0.51***			
		(0.04)	(0.03)	(0.36)				(0.07)				(0.25)				(0.05)		(0.08)			
ma3		0.10**	−0.40***									0.20**									
		(0.04)	(0.03)									(0.08)									
ma4												−0.04									
												(0.13)									
omega	0.00	0.00***	−3.69*	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00	0.00***	0.00**	0.00***	0.00***	0.00***	0.00	0.00***	0.00	0.00***	0.00***	0.00
	(0.00)	(0.00)	(2.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
alpha1	0.01	0.51***	0.04	0.44***	0.23**	0.32**	0.09	0.17*	0.10**	0.32	0.61**	0.28**	0.31**	0.28**	0.32***	0.04	0.34***	0.14	0.93***	0.02***	0.15*
	(0.01)	(0.13)	(0.10)	(0.14)	(0.09)	(0.15)	(0.06)	(0.10)	(0.05)	(0.20)	(0.25)	(0.14)	(0.14)	(0.12)	(0.12)	(0.04)	(0.10)	(0.10)	(0.21)	(0.01)	(0.08)
beta1	0.99***		0.43	0.19						0.44		0.46**				0.91***	0.41***	0.80***	0.07		0.65***
	(0.01)		(0.31)	(0.14)						(0.27)		(0.19)				(0.10)	(0.13)	(0.12)	(0.06)		(0.22)
delta																				3.50***	
																				(0.05)	
gamma1			0.44***								0.85*									0.94***	
			(0.14)								(0.52)									(0.01)	
skew	0.95***	0.92***		0.77***							0.84***										1.54***
	(0.11)	(0.10)		(0.12)							(0.08)										(0.29)
shape			7.29**	6.00**	8.15**	7.06***	8.53**	5.27***	6.89***	4.63***	3.02***	7.14***	7.90***	5.37***	94.24***	5.61***	6.69***		3.51***	7.30***	6.37***
			(3.09)	(2.46)	(3.42)	(2.50)	(3.63)	(1.37)	(2.53)	(1.08)	(0.45)	(2.64)	(2.98)	(2.01)	(22.73)	(1.82)	(2.39)		(0.66)	(0.63)	(2.12)
Num. obs.	291.00	291.00	291.00	291.00	291.00	291.00	291.00	291.00	291.00	291.00	291.00	291.00	291.00	291.00	291.00	291.00	291.00	291.00	291.00	291.00	291.00

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

Table A.7: US data sources

Variable	Description	Source
volSP500	Combined series of the annualized monthly volatility of the daily SP500 returns (until 12/1985) and the implied volatility index from SP500 option (VXO) (from 1/1986 on)	Yahoo Finance
d.rec.US	Dummy variable based on the US Business Cycle Expansions and Contractions data provided by The National Bureau of Economic Research (NBER)	Federal Reserve Economic data
IntR	Real Federal Funds Rate; effective Federal Funds Rate less realized inflation measured by year-on-year growth rate of CPI	Federal Reserve Economic data
SP500	Log Inflation adjusted (CPI) SP500	Yahoo Finance, Federal Reserve Economic data
oil	Log Inflation adjusted (CPI) Spot Crude Oil Price: West Texas Intermediate (WTI)	Federal Reserve Economic data
BCI	Business Confidence Index	OECD

Table A.8: German data sources

Variable	Description	Source
volDAX	Combined series of the annualized monthly volatility of the daily DAX returns (until 10/2005) and the implied volatility index from DAX option (VDAX) (from 11/2005 on)	Yahoo Finance
d.rec.GE	Dummy variable for recessions; derived from the growth rates of seasonally adjusted real Gross Domestic Product	Federal Statistical Office
IntR	Combined series of German Discount rate until 12/1998 and ECB policy Rate less realized inflation measured by year-on-year growth rate of CPI	Federal Reserve Economic data, Interntional Monetary Fund, OECD
DAX	Log Inflation adjusted (CPI) German stock market index (DAX)	Yahoo Finance, OECD
oil	Log Inflation adjusted (CPI) Spot Crude Oil Price: West Texas Intermediate (WTI)	Federal Reserve Economic data
BCI	Business Confidence Index	OECD

Appendix B

Appendix Chapter 4

Table B.1: Month-specific effects of Rain

	Dependent variable:	
	Nominal Retail turnover (dl)	
	(1)	(2)
Constant	0.041 (0.093)	−0.023 (0.100)
Δc_{t-1}	−0.756*** (0.075)	−0.750*** (0.080)
Δc_{t-2}	−0.534*** (0.085)	−0.472*** (0.088)
Δc_{t-3}	−0.217*** (0.075)	−0.188** (0.077)
R_t	0.452*** (0.119)	0.409*** (0.130)
$Rain_t^{Jan}$		0.014* (0.008)
$Rain_t^{Feb}$		−0.003 (0.010)
$Rain_t^{Mar}$		−0.012 (0.009)
$Rain_t^{Apr}$		−0.009 (0.008)
$Rain_t^{May}$		−0.003 (0.007)
$Rain_t^{Jun}$		−0.002 (0.008)
$Rain_t^{Jul}$		−0.007 (0.007)
$Rain_t^{Aug}$		0.015** (0.006)
$Rain_t^{Sep}$		−0.014 (0.010)
$Rain_t^{Oct}$		0.002 (0.008)
$Rain_t^{Nov}$		0.001 (0.005)
$Rain_t^{Dec}$		−0.004 (0.006)
$Rain_{t-1}^{Jan}$		0.008 (0.007)
$Rain_{t-1}^{Feb}$		−0.009 (0.008)
$Rain_{t-1}^{Mar}$		−0.005 (0.011)
$Rain_{t-1}^{Apr}$		0.003 (0.009)
$Rain_{t-1}^{May}$		0.014** (0.007)
$Rain_{t-1}^{Jun}$		−0.008 (0.007)
$Rain_{t-1}^{Jul}$		−0.0003 (0.008)
$Rain_{t-1}^{Aug}$		0.003 (0.006)
$Rain_{t-1}^{Sep}$		−0.002 (0.006)
$Rain_{t-1}^{Oct}$		−0.017* (0.010)
$Rain_{t-1}^{Nov}$		−0.003 (0.008)
$Rain_{t-1}^{Dec}$		0.006 (0.005)
Observations	176	176
R^2	0.389	0.495
Adjusted R^2	0.374	0.398
Residual Std. Error	1.190 (df = 171)	1.170 (df = 147)
F Statistic	27.200*** (df = 4; 171)	5.140*** (df = 28; 147)

Note: s.e. in parentheses — *p<0.1; **p<0.05; ***p<0.01

Table B.2: Month-specific effects of Sunshine

	Dependent variable:	
	Nominal Retail turnover (dl)	
	(1)	(2)
Constant	0.041 (0.093)	0.003 (0.096)
Δc_{t-1}	-0.756^{***} (0.075)	-0.693^{***} (0.081)
Δc_{t-2}	-0.534^{***} (0.085)	-0.472^{***} (0.087)
Δc_{t-3}	-0.217^{***} (0.075)	-0.178^{**} (0.076)
R_t	0.452^{***} (0.119)	0.447^{***} (0.128)
Sun_t^{Jan}		-0.001 (0.018)
Sun_t^{Feb}		0.003 (0.011)
Sun_t^{Mar}		0.018^{**} (0.007)
Sun_t^{Apr}		0.011^{*} (0.006)
Sun_t^{May}		-0.005 (0.008)
Sun_t^{Jun}		-0.005 (0.007)
Sun_t^{Jul}		0.013 (0.008)
Sun_t^{Aug}		-0.024^{***} (0.007)
Sun_t^{Sep}		0.002 (0.013)
Sun_t^{Oct}		-0.013 (0.015)
Sun_t^{Nov}		-0.003 (0.016)
Sun_t^{Dec}		0.0001 (0.011)
Sun_{t-1}^{Jan}		0.008 (0.012)
Sun_{t-1}^{Feb}		0.021 (0.018)
Sun_{t-1}^{Mar}		-0.008 (0.012)
Sun_{t-1}^{Apr}		-0.004 (0.007)
Sun_{t-1}^{May}		-0.014^{**} (0.006)
Sun_{t-1}^{Jun}		0.002 (0.008)
Sun_{t-1}^{Jul}		-0.003 (0.007)
Sun_{t-1}^{Aug}		0.009 (0.008)
Sun_{t-1}^{Sep}		0.011 (0.008)
Sun_{t-1}^{Oct}		0.026^{*} (0.014)
Sun_{t-1}^{Nov}		-0.012 (0.016)
Sun_{t-1}^{Dec}		0.006 (0.014)
Observations	176	176
R^2	0.389	0.522
Adjusted R^2	0.374	0.431
Residual Std. Error	1.190 (df = 171)	1.140 (df = 147)
F Statistic	27.200^{***} (df = 4; 171)	5.730^{***} (df = 28; 147)

Note: s.e. in parentheses $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table B.3: Month-specific effects of Temperature - Lasso specification

	Dependent variable:
	Nominal Retail turnover (dl)
Constant	0.030 (0.092)
Δc_{t-1}	-0.602^{***} (0.064)
Δc_{t-2}	-0.338^{***} (0.064)
R_t	0.430^{***} (0.104)
$D^{Mar}Temp_t$	0.591^{***} (0.195)
$D^{Apr}Temp_t$	0.220 (0.135)
$D^{Jun}Temp_t$	-0.250^{*} (0.128)
$D^{Aug}Temp_t$	-0.628^{***} (0.156)
$D^{Sep}Temp_t$	-0.356^{*} (0.181)
$D^{Oct}Temp_t$	-0.646^{***} (0.211)
$D^{Nov}Temp_t$	-0.249 (0.161)
$D^{May}Temp_{t-1}$	-0.380^{***} (0.134)
$D^{Aug}Temp_{t-1}$	0.252 (0.157)
$D^{Sep}Temp_{t-1}$	0.531^{***} (0.159)
$D^{Oct}Temp_{t-1}$	0.465^{**} (0.197)
$D^{Nov}Temp_{t-1}$	0.178 (0.201)
$D^{Dec}Temp_{t-1}$	0.493^{***} (0.159)
Observations	176
R^2	0.577
Adjusted R^2	0.535
Residual Std. Error	1.030 (df = 159)
F Statistic	13.600^{***} (df = 16; 159)
Note: s.e. in parentheses	*p<0.1; **p<0.05; ***p<0.01

Table B.4: Month-specific effects of Temperature - Real retail sales

	Dependent variable:	
	Real Retail turnover (dl)	
	(1)	(2)
Constant	0.284^{***} (0.097)	0.245^{***} (0.084)
Δc_{t-1}	-0.741^{***} (0.076)	-0.596^{***} (0.070)
Δc_{t-2}	-0.493^{***} (0.088)	-0.372^{***} (0.079)
Δc_{t-3}	-0.171^{**} (0.078)	-0.099 (0.070)
$IntR_t$	0.258^{*} (0.152)	0.282^{**} (0.133)
$D^{Jan}Temp_t$		0.175 (0.128)
$D^{Feb}Temp_t$		0.095 (0.091)
$D^{Mar}Temp_t$		0.367^{**} (0.150)
$D^{Apr}Temp_t$		0.334^{***} (0.102)
$D^{May}Temp_t$		-0.059 (0.145)
$D^{Jun}Temp_t$		-0.146 (0.097)
$D^{Jul}Temp_t$		-0.069 (0.116)
$D^{Aug}Temp_t$		-0.559^{***} (0.117)
$D^{Sep}Temp_t$		-0.366^{***} (0.135)
$D^{Oct}Temp_t$		-0.413^{***} (0.149)
$D^{Nov}Temp_t$		-0.413^{***} (0.120)
$D^{Dec}Temp_t$		0.108 (0.135)
Observations	173	173
R^2	0.368	0.561
Adjusted R^2	0.352	0.516
Residual Std. Error	1.230 (df = 168)	1.070 (df = 156)
F Statistic	24.400^{***} (df = 4; 168)	12.500^{***} (df = 16; 156)
Note: s.e. in parentheses		*p<0.1; **p<0.05; ***p<0.01

Table B.5: Month-specific effects of Temperature - year-on-year growth rates

	Dependent variable:	
	Nominal Retail turnover (yoy)	
	(1)	(2)
Constant	0.350** (0.144)	0.305** (0.140)
Δc_{t-1}	0.108 (0.077)	0.211*** (0.080)
Δc_{t-2}	0.146* (0.076)	0.113 (0.078)
Δc_{t-3}	0.229*** (0.076)	0.224*** (0.076)
Δc_{t-12}	−0.185*** (0.063)	−0.170*** (0.063)
R_t	1.370*** (0.255)	1.220*** (0.250)
$D^{Jan}Temp_t$		0.016 (0.184)
$D^{Feb}Temp_t$		0.141 (0.136)
$D^{Mar}Temp_t$		0.602*** (0.223)
$D^{Apr}Temp_t$		0.301** (0.142)
$D^{May}Temp_t$		0.201 (0.214)
$D^{Jun}Temp_t$		0.163 (0.252)
$D^{Jul}Temp_t$		−0.002 (0.172)
$D^{Aug}Temp_t$		−0.189 (0.213)
$D^{Sep}Temp_t$		−0.496** (0.200)
$D^{Oct}Temp_t$		−0.419* (0.235)
$D^{Nov}Temp_t$		−0.343* (0.185)
$D^{Dec}Temp_t$		−0.241 (0.204)
Observations	156	156
R^2	0.589	0.648
Adjusted R^2	0.575	0.605
Residual Std. Error	1.550 (df = 150)	1.500 (df = 138)
F Statistic	43.000*** (df = 5; 150)	14.900*** (df = 17; 138)
Note: s.e. in parentheses		*p<0.1; **p<0.05; ***p<0.01

Table B.6: Month-specific effects of Temperature with unemployment rate

	Dependent variable:	
	Nominal Retail turnover (dl)	
	(1)	(2)
Constant	0.044 (0.082)	−0.660 (0.714)
Δc_{t-1}	−0.605*** (0.069)	−0.608*** (0.069)
Δc_{t-2}	−0.408*** (0.078)	−0.410*** (0.078)
Δc_{t-3}	−0.127* (0.068)	−0.128* (0.068)
R_t	0.384*** (0.104)	0.450*** (0.120)
Δu_t	−1.060 (1.570)	
u_t		0.228 (0.233)
$D^{Jan}Temp_t$	0.181 (0.124)	0.179 (0.124)
$D^{Feb}Temp_t$	0.103 (0.088)	0.104 (0.088)
$D^{Mar}Temp_t$	0.352** (0.143)	0.351** (0.143)
$D^{Apr}Temp_t$	0.298*** (0.097)	0.299*** (0.097)
$D^{May}Temp_t$	−0.054 (0.138)	−0.042 (0.138)
$D^{Jun}Temp_t$	−0.165* (0.094)	−0.166* (0.094)
$D^{Jul}Temp_t$	−0.093 (0.113)	−0.093 (0.113)
$D^{Aug}Temp_t$	−0.564*** (0.114)	−0.560*** (0.114)
$D^{Sep}Temp_t$	−0.334** (0.131)	−0.336** (0.131)
$D^{Oct}Temp_t$	−0.423*** (0.143)	−0.423*** (0.143)
$D^{Nov}Temp_t$	−0.376*** (0.116)	−0.374*** (0.116)
$D^{Dec}Temp_t$	0.091 (0.131)	0.091 (0.131)
Observations	176	176
R^2	0.574	0.575
Adjusted R^2	0.528	0.530
Residual Std. Error (df = 158)	1.030	1.030
F Statistic (df = 17; 158)	12.500***	12.600***
Note: s.e. in parentheses		*p<0.1; **p<0.05; ***p<0.01

Table B.7: Month-specific effects of Temperature with exchange rate

	Dependent variable:	
	Nominal Retail turnover (dl)	
	(1)	(2)
Constant	0.053 (0.081)	−0.579 (1.090)
Δc_{t-1}	-0.591^{***} (0.069)	-0.604^{***} (0.069)
Δc_{t-2}	-0.395^{***} (0.077)	-0.407^{***} (0.078)
Δc_{t-3}	-0.130^{*} (0.068)	-0.125^{*} (0.068)
R_t	0.370^{***} (0.104)	0.294 (0.198)
$\Delta ExcR_t$	0.087 (0.055)	
$ExcR_t$		0.457 (0.806)
$D^{Jan}Temp_t$	0.165 (0.124)	0.181 (0.124)
$D^{Feb}Temp_t$	0.094 (0.087)	0.104 (0.088)
$D^{Mar}Temp_t$	0.362^{**} (0.142)	0.353^{**} (0.143)
$D^{Apr}Temp_t$	0.288^{***} (0.097)	0.300^{***} (0.097)
$D^{May}Temp_t$	−0.041 (0.137)	−0.045 (0.138)
$D^{Jun}Temp_t$	-0.166^{*} (0.093)	-0.166^{*} (0.094)
$D^{Jul}Temp_t$	−0.078 (0.113)	−0.093 (0.113)
$D^{Aug}Temp_t$	-0.545^{***} (0.114)	-0.561^{***} (0.114)
$D^{Sep}Temp_t$	-0.352^{***} (0.130)	-0.336^{**} (0.131)
$D^{Oct}Temp_t$	-0.434^{***} (0.143)	-0.424^{***} (0.143)
$D^{Nov}Temp_t$	-0.388^{***} (0.115)	-0.375^{***} (0.116)
$D^{Dec}Temp_t$	0.071 (0.131)	0.096 (0.131)
Observations	176	176
R^2	0.580	0.574
Adjusted R^2	0.534	0.528
Residual Std. Error (df = 158)	1.030	1.030
F Statistic (df = 17; 158)	12.800^{***}	12.500^{***}
Note: s.e. in parentheses		$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Appendix C

Appendix Chapter 5

Table C.1: Estimation results of α

	Dependent variable:
	α
ln(n)	-0.322^{***}
	(0.004)
Observations	5,102
R^2	0.587
Adjusted R^2	0.587
Note: s.e. in parentheses	*p<0.1; **p<0.05; ***p<0.01

Table C.2: Estimation results of γ

	Dependent variable:
	γ
Constant	2.625^{***}
	(0.154)
Observations	242
R^2	0.000
Adjusted R^2	0.000
Note: s.e. in parentheses	*p<0.1; **p<0.05; ***p<0.01

Curriculum Vitae

Personal Details

- Name: Anna Pauliina Sandqvist
- Date of Birth: July 1, 1985
- Place of Birth: Oulu, Finland
- Citizenship: Finnish

Current Position

- Research Assistant and Expert for Business Tendency Surveys and private consumption, KOF Swiss Economic Institute, ETH Zurich, Switzerland (since 7/2012)

Education

- Ph.D. student in Economics, ETH Zurich, Switzerland (since 5/2013)
  - *Courses at ETH:*
    * PhD Course in Panel Econometrics - Prof. Dr. Sturm, Prof. Dr. Pfaffermayr - ETH Zurich - Spring 2013
    * PhD Course in Program Evaluation Econometrics - Prof. Dr. Sturm, Dr. Basten, Dr. Moser - ETH Zurich - Spring 2013

    * Multivariate Statistics - Prof. Dr. Maathuis - ETH Zurich - Spring 2014
    * Programming with R for Reproducible Research - Dr. Mächler - Spring 2014
    * Development Economics II - Dr. Egger, Dr. Crole-Rees - ETH Zurich - Autumn 2014
    * Time Series Analysis - Prof. Dr. Meinshausen - Autumn 2015
  - *Courses outside ETH:*
    * Econometrics for Research Students Part 2 - Prof. Dr. Winkelmann - University of Zurich - Spring 2013
- Dipl. oec. in Economics, University of Hamburg, Germany, 2004-2009

KOF Dissertation Series

No 1 **Causes and Consequences of Democracy and Terror**
Martin Gassebner

No 2 **The Impact of News on Expectation Formation in Monetary Policy, Business Cycles and Inflation Forecasts**
Michael J. Lamla

No 3 **Micro Price Adjustments and Macro Inflation Dynamics**
Sarah M. Lein

No 4 **Essays on Inflation Expectation Formation**
Thomas Maag

No 5 **Essays in Political Economy and Fiscal Policy**
Frank S. Somogyi

No 6 **Modern Statistical Methods Applied to Economic Time Series**
Simone G. Elmer

No 7 **Analyzing Capital Investment and Unemployment Using Modern Statistical Methods**
Rolf M. Schenker

No 8 **Capacity Utilisation and Macro Dynamics from a Micro Perspective**
Eva M. Köberl

No 9 **Essays on the Production and Measurement of Knowledge Capital**
Thomas Bolli

No 10 **Deviations from Rationality – Essays on Inflation Perceptions and Expectations**
Lena C. R. Dräger

No 11 **Determinants of Training and the Performance of New Firms**
Tobias Stucki

No 12 **Measuring Media Sentiment – Essays on Its Impact on the Economy and the Financial Markets**
Matthias W. Uhl

No 13 **Rigidities, Institutions and the Macroeconomy**
Heiner Mikosch

No 14 **Applying Modern Econometric Methods to Swiss Unemployment and Consumption**
Martin Straub

No 15 **Determinants of Cross-Border Activity – Essays on Trade and Entrepreneurship**
Andrea T. Lassmann

No 16 **Three Essays on International Trade Policy and Welfare**
Sergey Nigai

No 17 **The Generation and Diffusion of Energy Efficient Technologies in Switzerland: An Empirical Study**
Marius Ley

No 18 **Four Essays on the Economics of Crime**
Marie Poprawe

No 19 **Examining the Special Case – Special Trends and Events at the Swiss Labor Market**
Michael Siegenthaler

No 20 **Empirical Studies on Research Productivity, Labour Mobility, and the Swiss Hospitality Industry**
Jörg Schläpfer

No 21 **The Globalization of Goods and Capital Flows**
Benedikt Rydzek

No 22 **Taxation and Firms' Investment Decisions**
Simon Bösenberg

No 23 **Essays on Applied Nonparametric Econometrics**
Sabrina Maria Dorn

No 24 **Four Essays on the Estimation of Time-Invariant Variables in Panel Data Models**
Michaela Kesina

No 25 **Survey Based Research in Economics – Essays on Methodology, Economic Applications and Long Term Processing of Economic Survey Data**
Matthias Bannert

No 26 **News and the Economy – How to Measure Economic Trends by Using Media-Based Data**
David Iselin

No 27 **Labor Income Taxation in a Globalizing World: 1980–2012**
Nora Margot Strecker

No 28 **On the Heterogeneous Effects of Trade and Fiscal Policy**
Katharina Margarete Klara Erhardt

No 29 **Forecasting at the Current Edge – Revisions and Mixed-Frequency Data**
Stefan Gerald Neuwirth

No 30 **Employment Protection and Labor Market Outcomes**
Johanna Mirka Kemper

No 31 **Broader Effects of Migration**
Jan Ruffner

No 32 **The Relationship Between Knowledge Absorption and Innovation Outcomes**
Florian Seliger

No 33 **Four Essays in Wine Economics: An Empirical Approach with Swiss Panel Data**
Alexandre Mondoux

No 34 **The Macroeconomics of Uncertainty**
Andreas Dibiasi

No 35 **Essays on Business Cycles and Sectoral Dynamics**
Anna Pauliina Sandqvist

KOF

ETH Zurich
KOF Swiss Economic Institute
LEE G 116
Leonhardstrasse 21
8092 Zurich, Switzerland

Phone +41 44 632 42 39
Fax +41 44 632 12 18
kof@kof.ethz.ch
www.kof.ethz.ch
#KOFETH