tinbergen institute

Unraveling Dimensions:
Commodity Futures Curves and Equity Liquidity

Dennis Karstanje

Erasmus Universiteit Rotterdam

(This page is intentionally left blank)

Unraveling Dimensions:

Commodity futures curves and equity liquidity

Unraveling Dimensions:
Commodity futures curves and equity liquidity

Dimensies uitgezocht:
Grondstofcurven en de liquiditeit van aandelen

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

prof.dr. H.A.P. Pols

and in accordance with the decision of the Doctorate Board.

The public defense shall be held on

Thursday, March 26, 2015 at 13:30 hours

by

DENNIS KARSTANJE
born in Vlissingen, The Netherlands

Erasmus University Rotterdam

Doctorate Committee

Promotor:	Prof.dr. D.J.C. van Dijk
Other members:	Dr. R. Huisman
	Prof.dr. A. Lunde
	Prof.dr. M.A. van Dijk
Copromotors:	Dr. W.W. Tham
	Dr. M. van der Wel

ISBN: 978 90 361 0425 8

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. 606 of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

Acknowledgments

Being a PhD student has had a profound impact on my life. It allowed me to learn more about econometrics, finance, commodities, programming and many other things. It made it possible to travel around the world to present my work, meet new people and experience other cultures. In short, it enriched many dimensions of my life.

I owe a large debt of gratitude to my promoter, Dick van Dijk, and my co-promoters, Michel van der Wel and Wing Wah Tham. Dick, you were always available to give me guidance, but at the same time you gave me the freedom to explore and develop my own ideas. It was a privilege to work with you; you are a true inspiration. Michel, you are the only person who is coauthor on two of my chapters. Thank you for all your input and advice. I will never forget your positive vibe, reflected by your motto: “There are no problems in life, only challenges”. Wing Wah, without you as a supervisor during my FCS project I probably would never have started as a PhD student. I enjoyed our intellectual chats and your constructive criticism pushed my research to a higher level.

More than four years ago, Wilma de Groot and Weili Zhou gave me the chance to do an internship at Robeco. Both of you, lighted my interest in commodities. I enjoy(ed) working together and I am proud of Chapter 3, which started out as my master thesis but made its way into a peer-reviewed journal.

I would like to thank Elvira Sojli, coauthor on my second chapter, for the great collaboration. During the first two years of my PhD it felt like I had an additional supervisor. Thank you for sharing your experience and all your academic tips and tricks.

Further, I am indebted to Ronald Huisman, Asger Lunde and Mathijs van Dijk for taking place in my doctoral committee, for reading my thesis thoroughly and for their helpful comments. I am also grateful to Marta Szymanowska and Martin Martens for taking place in the larger committee.

On January 3, 2011 I entered the Erasmus university as a PhD student. Since the academic year was already halfway, the Tinbergen courses I started with were at times quite overwhelming. Thanks to Maarten and Inez, I quickly felt at ease. I enjoyed working together on assignments and the support of both of you in these first “rough” months. While not taking courses, I spent a lot of time on the ninth floor. Thank you, Kyle and Irena, with whom I shared an office. The both of you were opposites but I liked the time we spent together. After the move to the eighth floor, I shared an office with Anne. I will never forget your stories, anecdotes and humor. We had lots of fun both in- and outside the office and I consider you as a true friend.

The list of fellow PhDs with whom I share great memories is long. Knowing that I am forgetting people, I want to thank: Eran, for the great time we had in Rotterdam and at several conferences. Martin, for dropping by in my office from day one and many days thereafter. Bart, Bert, Bruno, Didier, Francine, Koen, Myrthe, Sander and Tom for all joint lunches, coffee breaks, many Vrimibo’s and of course Pukkelpop. Barbara, Justinas, Mehtap, Niels, Sait, Sanne, Saskia, Victor, and many others for the great time I had during my PhD.

Furthermore, I would like to thank all colleagues at the eleventh floor, in particular Bart, Erik, Marcin, Tommi and Twan; Suzanne, who is always in for a chat and some coffee; all colleagues at the EDSC Datateam: Paul, Judith, Ronald, Sharon, Stefan and Sven; everyone at the Tinbergen Institute and Econometric Institute for making my PhD research possible. Thank you Anneke, Barbara, Carien, Carine, Carolien, Ine, Judith, Marianne, Marjon and Ursula; and all colleagues at Robeco for providing a fun and stimulating working environment both before and after my PhD.

A special thanks goes out to my paranimphs Bart and Bruno. We had lots of fun during the last year. Bart, I enjoyed the time we shared an office and all our chats and discussions fueled by either coffee or beer. Bruno, I like our shared interest in programming, technology, and many more things (not to forget our escapades with Lisa and Julia). Of course, this thesis would not be completed without all the rounds of mini table tennis I played with both of you.

Besides thanking all my colleagues I would like to thank my friends for their support and to made me realize that there is more to life than just research. You were there when I needed distraction the most. Thank you Andry, Janine, Marnix, Heidi, Stijn, Morindy, Christoph, Jorein, Marco, Stefan and Tigran.

Last, I thank my family for their love and support. In particular, Leny and Piet, for being an example and inspiration. Robin and Jamie, for their support and all the fun we had. Ronald and Anita, for your interest and encouragement. Ina, unfortunately you will never be able to read this thesis. I never saw you as my aunt but more as a second mom. Eline, you will always be my little sister of whom I am proud. Mom and dad, thanks for being the parents all one could wish for, for always supporting my decisions in life, for always being available whenever I need a listening ear, for your love.

At the end I would like to express appreciation to Esther. More than anyone else, you know the highs and lows I experienced during my PhD. Thank you for being there and sticking by my side. There are no words to convey how much I love you. We have been through a lot, but in your arms I feel safe.

Dennis Karstanje

Rotterdam, 2015

Contents

1 Introduction and outline **1**

1.1 Liquidity . . . 2

1.2 Commodity futures curves . . . 4

1.3 Outline . . . 6

2 Economic valuation of liquidity timing **9**

2.1 Introduction . . . 9

2.2 Methodology . . . 12

2.2.1 Forecasting liquidity and expected returns . . . 12

2.2.2 Asset allocation . . . 13

2.2.3 Evaluation . . . 14

2.3 Data . . . 17

2.3.1 Portfolio construction . . . 18

2.3.2 Liquidity variables . . . 19

2.3.3 Preliminary statistics . . . 22

2.4 Results . . . 26

2.4.1 Main results . . . 27

2.4.2 Control variables . . . 30

2.4.3 Cross-sectional predictors . . . 33

2.4.4 Risk adjusted returns . . . 34

2.5 Why Zeros performs best? . . . 36

2.5.1 Performance conditional on market returns . . . 36

2.5.2 Quality of predicted returns . . . 37

2.6 Robustness . . . 39

2.6.1 Benchmarks . . . 39
2.6.2 Sensitivity analysis . . . 40
2.6.3 Bias adjustment . . . 41
2.7 Conclusion . . . 42
2.A Robustness results . . . 43

3 Exploiting commodity momentum along the futures curves 49
3.1 Introduction . . . 49
3.2 Data . . . 54
3.3 Methodology . . . 60
3.3.1 Constructing momentum strategies . . . 60
3.3.2 Incorporating transaction costs . . . 62
3.4 Results . . . 64
3.4.1 Profitability of strategies including term-structure information . . . 64
3.4.2 Portfolio return regressions . . . 68
3.4.3 Double-sort on momentum and carry . . . 72
3.5 Liquidity analyses . . . 75
3.5.1 Implementation with futures contracts up to six months maturity . . . 75
3.5.2 Implementation on most liquid futures contracts . . . 77
3.5.3 Results with a one-day implementation lag . . . 81
3.5.4 Results since 2000 . . . 83
3.6 Conclusion . . . 85

4 Common factors in commodity futures curves 87
4.1 Introduction . . . 87
4.2 Methodology . . . 91
4.2.1 Model . . . 91
4.2.2 Estimation . . . 94
4.3 Data . . . 95
4.4 Individual commodity results . . . 99
4.4.1 Estimation results individual factors . . . 99
4.4.2 Comparison with the Schwartz (1997) three-factor model . . . 107
4.5 Joint model for commodity curves . . . 111

4.5.1 Commonality results 111
4.5.2 Importance of common factors 113
4.5.3 Factor dynamics 116
4.5.4 Economic interpretation of unobserved states 122
4.5.5 Time variation 127
4.5.6 Importance curve data 129
4.6 Conclusion 131
4.A State space representation 133
4.B Maturity bound 135
4.C Additional individual commodity results 138
4.D Additional Schwartz comparison results 142
4.E Macroeconomic data 144

Nederlandse samenvatting (Summary in Dutch) 151

Bibliography 155

Chapter 1

Introduction and outline

Open your mind to new ideas

Abbott (1884)

An inspiring and intriguing book on dimensions is "Flatland - A Romancy of Many Dimensions" by Edwin A. Abbott.[1] In this satirical novel, Abbott describes different dimensional worlds from the point of view of A. Square, who is living in the two-dimensional world called Flatland. The Square dreams about a visit to a one-dimensional world (Lineland) inhabited by Points. He attempts to convince them of the existence of a second dimension, but is unable to do so. When A. Square is visited by a Sphere (from Spaceland), he cannot comprehend this third dimension until he sees Spaceland for himself. After the Square's mind is opened to new dimensions, he tries to convince the Sphere of the theoretical possibility of the existence of a fourth (and fifth, and sixth, ...) spatial dimension, but the Sphere returns his student to Flatland in disgrace. The Square recognizes the identity of the ignorance of the inhabitants of Pointland and Lineland with his own (and the Sphere's) previous ignorance of the existence of higher dimensions.

The link between Flatland and this thesis is the concept of dimensions. Many dimensions exist in the world of financial markets, much more than we can comprehend as inhabitants of our three-dimensional world. As I will outline below, many dimensions in financial markets remain uninvestigated. The two that have the main focus in this thesis are the dimensions of liquidity in the equities market and the term-structure dimension in the commodities market.

[1] Although Flatland was not ignored when it was published, it did not obtain a great success. The book was discovered again after Albert Einstein's general theory of relativity was published and was also mentioned in Nature (Garnett, 1920).

Both are not yet explored to their full extent. Besides this direct link there is another link between the book, Flatland, and this thesis. The essential message of the book is that we should "open our mind to new ideas", which is very similar to the journey of a PhD student. We should explore fields that we have not done before and we should be receptive to things that baffle us at first thought.

The goal of this thesis is to present frameworks that shed more light on the mentioned dimensions and to use the new findings and insights in ways that are relevant from both an academic and practical point of view. The thesis consists of two parts. In the first part liquidity is at the center stage, while in the second part we focus on the term structure dimension in commodity futures.

1.1 Liquidity

While seemingly a simple concept, the exact meaning of liquidity is far from apparent. The fact that liquidity is unobserved and the existence and disagreement over various definitions, make it an elusive concept.[2] As a starting point we might agree that liquidity relates to the ability to buy and sell assets easily. Harris (2003) provides a more detailed definition, "liquidity is the ability to trade large size quickly, at low cost, when you want to trade." Various dimensions of liquidity appear in this definition, namely: speed, impact, cost and timing.[3] All these different dimension are related. For example, if you want to trade a large amount of assets you will either have a large impact on prices or it will take you more time (and costs) when you split your trades.

Besides the existence of various definitions and dimensions of liquidity, another element that is important in this thesis is the relation between liquidity and asset prices, more specifically the existence of a liquidity effect. According to standard asset pricing theory, assets with the same cash flow d should have the same price p. To study if and how liquidity is priced empirically, we can compare the prices of assets with the same stream of cash flows

[2] O'Hara (1995) draws on an analogy with pornography: "it is hard to define, but you know it when you see it."

[3] Other dimensions often mentioned are width, depth, breadth and resiliency. Depth is related to the quantities available in the order book, i.e. below (above) the current market price there is a large quantity available for sale (to buy). Breadth is related to the number of participants and their market power. In a resilient market, effects due to trading die out quickly.

but different liquidity values.[4] In contrast to standard asset pricing, liquidity-based asset pricing argues that if asset 1 is less liquid than asset 2, which we notate as $L_1 < L_2$, but both have the same cash flows $d_1 = d_2$, the price of asset 1 should be lower $p_1 < p_2$. In practice it is difficult to find two assets with the same stream of cash flows, hence we need to control for differences in other determinants. Furthermore, instead of looking at differences in price, we can also look at differences in assets' expected returns.

A major problem in estimating the effect of liquidity on asset prices or returns is how to measure liquidity, since there is hardly a single measure that captures all of its aspects (Amihud et al., 2005). Furthermore, studies of the effect of liquidity on expected stock returns use ex-post or realized returns, whose variance around the expected return is high. Consequently, researchers need a large amount of data to increase the power of their tests. Given that high-frequency data is only available from the beginning of the nineties, this poses a problem. Especially, when one wants to cover various market conditions, e.g. the entire business cycle, or when one is interested in a longer term effect. Researchers need then to find substitute measures of liquidity using lower frequency data. Examples of low-frequency liquidity measures are Roll (Roll, 1984), Zeros (Lesmond et al., 1999), Amihud *ILLIQ* (Amihud, 2002), and Effective Tick (Holden, 2009). Several papers provide a comparison between low and high-frequency liquidity measures (Lesmond et al., 1999; Lesmond, 2005; Hasbrouck, 2009; Goyenko et al., 2009).

One of the first papers that empirically investigates the liquidity effect is Amihud and Mendelson (1986). They find that expected asset return is an increasing function of illiquidity costs. Besides this cross-sectional effect, we can also investigate the effect of liquidity on returns over time. Amihud (2002), Jones (2002), and Baker and Stein (2004) show for the U.S. market that when liquidity is expected to be low, expected returns are higher. Bekaert, Harvey, and Lundblad (2007) find supporting evidence in emerging markets. Explanations for these findings are given by Amihud and Mendelson (1986) and Vayanos (1998) who argue that investors anticipate future transaction costs and discount assets with higher transaction costs more. Baker and Stein (2004) relate liquidity to irrational investors who under-react to information in order flow. These investors are restricted by short-sales constraints and only participate in the market when they overvalue the market relative to rational

[4]Different liquidity values can refer to either differences in liquidity levels or risks.

investors. Hence when the market is more liquid, it is overvalued and expected returns are lower.

Given the evidence for the existence of a liquidity effect, an investor could time liquidity. That is, if an investor can predict when the market will be liquid or illiquid, she can adjust her exposure before liquidity events occur. Cao, Chen, Liang, and Lo (2013) provide evidence that many hedge fund managers behave like liquidity timers, adjusting the market exposure of their portfolios based on equity-market liquidity. However, as discussed above, various definitions and dimensions of liquidity exist. Existing literature gives no guidance on empirical models and measures that an investor could use for liquidity timing. In the first part of this thesis we take this task upon us, and provide a solution.

1.2 Commodity futures curves

Commodities constitute the only spot markets which have existed nearly throughout the history of human kind (Geman, 2005). One of the earliest reports of futures markets is the market for tulips in The Netherlands.[5] These formal futures markets developed in 1636 and were the primary focus of trading. Thus, as a bet on the price of the bulbs on the settlement date, this market was not different in function from currently operating futures markets (Garber, 1989). Nowadays most commodities are traded in the U.S., where in the 18th and 19th centuries farmers started selling their crops at the time of planting in order to finance the production process.

Given the theme of this thesis, I would like to focus on the various dimensions in commodity futures markets. As their futures markets exist for several centuries, long data series over time are available. The properties of these time series have been and are still investigated in academic literature, examples are the existence of trends, cycles and pricing bubbles. The cross-sectional dimension, i.e. the comparison of different commodities, is also an area of research interest. Findings are interesting for planting decisions of farmers, hedging decisions of industrial corporations and asset allocation decisions of investors. Besides the time series and cross-sectional dimension, futures markets have an additional dimension corresponding to the different settlement or maturity dates on which futures contracts expire. This dimen-

[5] A futures contract is a contract between two parties to buy or sell an asset for a price agreed upon today (the futures price) with delivery and payment occurring at a future point, the maturity date.

sion is referred to as the term-structure or futures curve and is the dimension of interest in the second part of this thesis.

Commodity futures curves, which correspond to prices of futures contracts for different maturities, can be upward or downward sloping. A futures curve where prices are lower than the current spot price is called to be in backwardation, while a futures curve that is upward sloping is referred to as being in contango. An extensive literature focuses on explaining the shapes of commodity futures. Two prevailing theories are the theory of normal backwardation and the theory of storage. Keynes (1930) original theory of normal backwardation is based on the presumption that hedgers, commodity producers who want to secure a certain price level for future deliveries, hold on average a short position in the futures market. Since these market participants are willing to pay a risk premium in order to hedge their exposure to spot price positions, the price of a futures contract will be a downward biased estimator of future spot prices. Cootner (1960) generalized the theory by allowing hedgers to be net long, for example if there are many commodity consumers who want to hedge. Depending on the net position of hedgers, futures prices may carry either a positive or a negative risk premium. The theory of storage, originally proposed by Kaldor (1939), Working (1949), Brennan (1958), and Telser (1958), relates spot and futures contract prices to inventories and the stream of costs and benefits from holding the physical commodity. Central to the theory is the convenience, which is defined by Brennan and Schwartz (1985) as “the flow of services that accrues to an owner of the physical commodity but not to the owner of a contract for future delivery of the commodity”. When the benefits of holding the physical asset (net convenience yields) are higher than the financing costs (interest rates), the futures curve is in backwardation, whereas the futures curve is in contango when interest rates exceed the net convenience yield. Note that the arguments above hold for the relation between the spot price and all futures prices (regardless of their maturity). Hence, even when we are not interested in the spot price, it forces a structure on the futures curve.

Despite the extensive literature on modeling the shape of commodity futures curves, none of these papers incorporate these insights in trading strategies or commonality analysis. In the second part of this thesis we use the curve dimension to enhance existing strategies and extend modeling frameworks to deal with time-series, cross-sectional and term-structure dimensions.

1.3 Outline

We pursue the goals set forth above in three separate chapters, which are all self-contained and can be read independently. The next chapter deals with liquidity in equity markets while Chapters 3 and 4 focus on commodity futures curves.

Chapter 2 is based on Karstanje, Sojli, Tham, and van der Wel (2013). In this chapter, we focus on the dimensions of liquidity in the light of market timing. We conduct a horse-race of different liquidity proxies using dynamic asset allocation strategies to evaluate the short-horizon predictive ability of liquidity on monthly stock returns. We assess the economic value of the out-of-sample power of empirical models based on different liquidity measures and find three key results: liquidity timing leads to tangible economic gains; a risk-averse investor will pay a high performance fee to switch from a dynamic portfolio strategy based on various liquidity measures to one that conditions on the Zeros measure (Lesmond, Ogden, and Trzcinka, 1999); the Zeros measure outperforms other liquidity measures because of its robustness in extreme market conditions. These findings are stable over time and robust to controlling for existing market return predictors or considering risk-adjusted returns.

Chapter 3 is based on de Groot, Karstanje, and Zhou (2014). Here, we examine novel momentum strategies in commodities futures markets that incorporate term-structure information. We show that momentum strategies that invest in contracts on the futures curve with the largest expected roll-yield or the strongest momentum earn significantly higher risk-adjusted returns than a traditional momentum strategy, which only invests in the nearest contracts. Moreover, when incorporating conservative transaction costs we observe that our low-turnover momentum strategy more than doubles the net return compared to a traditional momentum strategy.

Chapter 4 is based on Karstanje, van der Wel, and van Dijk (2015). In this chapter, we examine the existence of common factors driving commodity futures curves. We adopt the framework of the dynamic Nelson-Siegel model, enabling us to examine not only commonality in price levels but also futures curve shapes, as characterized by their slope and curvature. Our empirical results based on 24 commodities over the period 1995-2012 demonstrate that the individual commodity futures curves are driven by common components. The commonality is mostly sector specific, which implies that commodities are a heterogeneous asset class. The common components in the level of the curve have become more important over time, coinciding with the financialization of the commodities market. The market-wide level component, which is common to all commodities, is related to economic output variables, exchange rates and hedging pressure. Factors driving the shape of the futures curve are related to inventory data (theory of storage), hedging pressure (theory of normal backwardation) and interest rates. The use of full curve data alters findings on commonality, compared to the use of only first-nearby contract data. The full curve commonality results give more insight in the market dynamics and can help in the construction of commodity futures portfolios and hedging decisions.

Overall, this thesis contributes to several strands of literature. First, we provide a framework to implement liquidity timing, contributing to the liquidity measure and liquidity-return literature. Using our framework we show that different dimensions of liquidity vary in their ability to predict expected future returns. Besides the direct results following from our approach, the proposed framework is also an alternative way to compare various low-frequency liquidity measures. By quantifying the "quality" of a liquidity measure in economic terms, it allows us to compare measures of different dimensions. Second, we show the added value of taking into account the term-structure dimension of commodity futures. We do this both in a practical setting and a more theoretical setting. In the practical setting we focus on an investment strategy and show that increasing the investment universe and incorporating curve information lead to better investment results. The theoretical setting focuses on the commonality across various commodities. Here, we propose a framework that can deal with three dimensions (time series, cross-sectional and term-structure) at once and use this framework to show that the inclusion of curve information is important. We show and explain how the incorporation of curve information leads to different results in terms of cross commodity comovement.

Chapter 2

Economic valuation of liquidity timing

Based on Karstanje, Sojli, Tham, and van der Wel (2013)

2.1 Introduction

There is ample evidence that liquidity, the ease with which financial assets can be bought and sold, is important in explaining variations in asset prices. When market liquidity is expected to be low expected returns are higher.[1,2] A smart investor can potentially time the market and adjust exposure before liquidity events occur, i.e. time liquidity. Cao, Chen, Liang, and Lo (2013) provide evidence that many hedge fund managers behave like liquidity timers, adjusting the market exposure of their portfolios based on equity-market liquidity. However there is no guidance on empirical models and measures that one could use for liquidity timing, and this paper addresses these issues.

The literature approximates the unobserved liquidity of a financial asset using various liquidity measures. A large number of proxies for liquidity exists because liquidity has multiple aspects (e.g. width, depth, immediacy, or resiliency). Examples of liquidity proxies are

[1]Amihud (2002), Jones (2002), and Baker and Stein (2004) show this for the U.S. market. Bekaert, Harvey, and Lundblad (2007) find supporting evidence in emerging markets.

[2]Amihud and Mendelson (1986) and Vayanos (1998) argue that investors anticipate future transaction costs and discount assets with higher transaction costs more. Baker and Stein (2004) relate liquidity to irrational investors who under-react to information in order flow. These investors are restricted by short-sales constraints and only participate in the market when they overvalue the market relative to rational investors. Hence when the market is more liquid, it is overvalued and expected returns are lower.

spread proxies, measures of price impact, and turnover.[3] However it is unclear what liquidity measure an investor should use for liquidity timing and how it should be implemented.

In this paper we examine which proxy a liquidity timer should use. We do so, by measuring the economic value of liquidity forecasts using different liquidity proxies, from the perspective of investors who engage in short-horizon asset allocation strategies. We focus on the economic valuation of liquidity because it is relevant from an investor's point of view. Moreover it allows us to compare the performance of different liquidity measures, which might be capturing different aspects of liquidity, under the same "unit".[4]

We consider the following five low-frequency liquidity measures for liquidity timing: illiquidity ratio (ILR) (Amihud, 2002), Roll (Roll, 1984), Effective Tick (Holden, 2009; Goyenko, Holden, and Trzcinka, 2009), Zeros (Lesmond, Ogden, and Trzcinka, 1999), and High-Low (Corwin and Schultz, 2012). Using these liquidity measures, we form conditional expectations about stock returns for the next period. Building on previous research (e.g. West, Edison, and Cho, 1993), we employ mean-variance analysis as a standard measure of portfolio performance and apply quadratic utility to examine and to compare the economic gains of the different measures. We use the Sharpe ratio (SR) and performance fee to evaluate the economic gains.[5] In addition, we also calculate the break-even transaction cost, which is the transaction cost that would remove any economic gain from a dynamic asset allocation strategy.

Based on NYSE-listed stocks for the period 1947-2008, we find evidence of economic value in liquidity timing. The Zeros measure outperforms the other measures: ILR, Roll, Effective Tick, and High-Low. The Zeros measure achieves a Sharpe ratio of 0.51, followed by the ILR with a Sharpe Ratio of 0.27. The SR of a buy and hold strategy over the same period is 0.28. A risk-averse investor with quadratic utility would pay an annual fee of more than 250 basis points to switch from the other liquidity proxies to condition on the Zeros

[3]For spread proxies see e.g. Roll (1984), Lesmond, Ogden, and Trzcinka (1999), Hasbrouck (2009), and Holden (2009); for price impact measures see e.g. Amihud, Mendelson, and Lauterbach (1997), Berkman and Eleswarapu (1998), Amihud (2002), and Pástor and Stambaugh (2003), and for turnover see Baker and Stein (2004).

[4]Other articles, e.g. Goyenko, Holden, and Trzcinka (2009) and Hasbrouck (2009), compare liquidity measures to a benchmark. Their set-up only allows to compare proxies that approximate the same aspect of liquidity, e.g. price impact or effective spread.

[5]The Sharpe ratio is the most common measure of performance evaluation employed in financial markets to assess the success or failure of active asset managers; it is calculated as the ratio of the average realized portfolio excess returns to their variability. The performance fee measures how much a risk-averse investor is willing to pay for switching from one strategy to another.

liquidity measure. The alpha of the Zeros strategy is 7.01% after controlling for exposure to the three Fama and French (1993) factors, the Carhart (1997) momentum factor, and the Pástor and Stambaugh (2003) liquidity factor. The results are not driven by correlations with other return predictors such as the dividend yield or the book-to-market ratio (Welch and Goyal, 2008). Furthermore, the outperformance is not specific to a particular period and is robust to different subsamples, weight restrictions, and target volatility and risk aversion parameters.

We document that the Zeros measure shows positive performance under all market conditions. Its returns remain very high throughout both bull and bear periods and its weights remain quite stable. Additionally, we show that the return predictions of the Zeros strategy are of good quality. We do this by restricting the weights in the asset allocation to be nonnegative. Jagannathan and Ma (2003) show that imposing nonnegativity restrictions in an asset allocation problem, reduces the estimation error in the return prediction parameters and gives similar effects as shrinking the return predictions. However, if the quality of the predictions is already good and cannot simply be improved by shrinkage, the strategy performance will deteriorate when restrictions are imposed. We find that weight restrictions lower the performance of the Zeros strategy, while they increase the performance of the other strategies.

This paper contributes to the literature on liquidity proxies comparison. Goyenko, Holden, and Trzcinka (2009) investigate how well low frequency liquidity measures approximate true transaction costs for market participants, which are measured by high-frequency benchmarks. They find that Effective Tick is the best low frequency measure for effective and realized spread, and ILR is the best measure for price impact. However, the best proxy for transaction costs is not necessarily the proxy that an investor should use for liquidity timing. In contrast, this paper investigates which measure can be used to time the market. Effective Tick shows no economic value, despite its ability to approximate high frequency transaction costs well and the Zeros measure is the most relevant for liquidity timing.

This paper contributes also to the literature on portfolio allocation. West, Edison, and Cho (1993) use the mean-variance and quadratic utility setting to rank exchange rate volatility models based on utility gains. Fleming, Kirby, and Ostdiek (2001) investigate volatility timing in equity markets. Della-Corte, Sarno, and Thornton (2008) and Della-Corte, Sarno, and Tsiakas (2009) apply the approach to short-term interest rates and predictability in the foreign exchange market. Thornton and Valente (2012) investigate the economic value of long-term forward interest rate information to predict bond returns. Differently from these papers, we evaluate the economic value of liquidity timing in equity markets.

2.2 Methodology

We examine whether liquidity timing leads to economic benefits and which liquidity proxy should be used, following three steps. First, we form conditional expectations of returns based on different liquidity measures. Second, we construct dynamically rebalanced mean-variance portfolios based on these return predictions. Third, we evaluate the performance of these strategies. In this section we focus on the methodology, while implementation details are presented when discussing the results.

2.2.1 Forecasting liquidity and expected returns

We start by modeling liquidity in order to estimate expected liquidity in the next period. Following Amihud (2002), Acharya and Pedersen (2005), and Bekaert, Harvey, and Lundblad (2007) we use autoregressive models to capture the autocorrelation in the liquidity series:

$$LIQ_{k,t} = \phi_0 + \sum_{i=1}^{p} \phi_i LIQ_{k,t-i} + \eta_{k,t}, \tag{2.1}$$

where $LIQ_{k,t}$ is the liquidity of asset k at time t, and p is the order of the autoregressive model. Iterating forward Equation (2.1), liquidity predictions for the next period are given by $E_t\left[LIQ_{k,t+1}\right] = \phi_{0,t} + \sum_{i=1}^{p} \phi_{i,t} LIQ_{k,t-i}$. Adding expected liquidity in a model for conditional expected excess returns that is solely driven by liquidity, gives:

$$\begin{aligned} E_t\left[r_{k,t+1} - r_{f,t}\right] &= \delta_0 + \delta_1 E_t\left[LIQ_{k,t+1}\right] \\ &= \delta_{0,t} + \delta_{1,t}\left(\phi_{0,t} + \sum_{i=1}^{p} \phi_{i,t} LIQ_{k,t-i}\right) \\ &= \beta_{0,t} + \sum_{i=1}^{p} \beta_{i,t} LIQ_{k,t-i}, \end{aligned} \tag{2.2}$$

where $\beta_{0,t} = \delta_0 + \delta_1 \phi_{0,t}$ and $\beta_{i,t} = \delta_1 \phi_i$. We only need estimates for the β-parameters and do not estimate Equation (2.1), because we are interested in return predictions generated by Equation (2.2). The coefficients $\beta_{0,t}$ and $\beta_{i,t}$ are allowed to vary over time and are estimated using a rolling window of length L. If liquidity is beneficial for forecasting expected returns, it can be used in a 'liquidity timing' strategy. We estimate the parameters in Equation (2.2) using a window length of 10 years ($L = 120$ monthly observations). To minimize the effect of possible structural breaks on the results, Pesaran and Pick (2011) suggest to average predictions generated using different rolling window lengths. We take the average of three different predictions based on a window length of 5, 10, and 20 years ($L = 60, 120$, and 240 monthly observations).[6]

To allow for a long enough sample to cover the longest moving window of 20 years, the first return prediction is made for January 1967. For the 10 year moving window, we estimate the regression in Equation (2.2) using data from January 1957 to December 1966. Using the estimated coefficients we make a forecast for next month, January 1967. Then we shift the window one period ahead. Thus the second estimation window runs from February 1957 to January 1967, and we make a prediction for February 1967. This procedure is repeated for all months $t =$ Jan 1967, Feb 1967, ..., Dec 2008 and all assets $k = 1, 2, \ldots, K$, for each liquidity measure. For the 5 year moving window, the first window is January 1962 to December 1966 and for the 20 years, the first window is January 1947 to December 1966.

2.2.2 Asset allocation

Following the literature, we use mean-variance dynamic trading strategies to assess the economic value of liquidity timing. An investor invests every month in the K risky assets and one riskless U.S. Treasury bill $(r_{f,t})$. She chooses the weights to invest in each risky asset by constructing a dynamically re-balanced portfolio that maximizes the conditional expected

[6]One could alternatively use a non-linear specification to forecast returns. However, we choose to use a linear over a non-linear specification throughout all of our analysis for two reasons. First, although a correctly specified non-linear model may fit the conditional expectation function more closely than a linear model, a misspecified non-linear model may perform worse. OLS provides a robust approach as the best linear estimator for the non-linear relation. Second, Clements et al. (2004) show that the forecasting performance of non-linear models is not particularly good compared to linear models. They conclude that the problem may be that non-linear models are not mimicking reality any better than simpler linear approximations.

return subject to a target conditional volatility. Her optimization problem is given by

$$\max_{w_t} \left\{ r_{s,t+1|t} = w_t' r_{k,t+1|t} + (1 - w_t'\mathbf{1}) r_{f,t} \right\} \tag{2.3}$$
$$\text{s.t. } (\sigma_s^*)^2 = w_t' \Sigma_{t+1|t} w_t,$$

where $r_{s,t+1|t}$ is the conditional expected return of strategy s, w_t is the vector of weights of the risky assets, $r_{k,t+1|t}$ is the vector of conditional risky asset return predictions, σ_s^* is the target level of risk for the strategy, and $\Sigma_{t+1|t}$ is the variance-covariance matrix of the risky assets. $\Sigma_{t+1|t}$ is estimated recursively as the investor updates return predictions and dynamically balances her portfolio every month.[7] The solution to this maximization problem yields the risky asset investment weights:

$$w_t = \frac{\sigma_s^*}{\sqrt{Q_t}} \Sigma_{t+1|t}^{-1} \left(r_{k,t+1|t} - \mathbf{1} r_{f,t} \right),$$

where $Q_t = \left(r_{k,t+1|t} - \mathbf{1} r_{f,t} \right)' \Sigma_{t+1|t}^{-1} \left(r_{k,t+1|t} - \mathbf{1} r_{f,t} \right)$ and $r_{k,t+1|t} - \mathbf{1} r_{f,t}$ is the conditional excess return. The weight invested in the risk free asset is $1 - w_t'\mathbf{1}$. The covariance matrix is estimated by the sample covariance matrix over a 10 year rolling window, thus, the covariance matrix is time-varying.[8]

2.2.3 Evaluation

We employ mean-variance analysis as a standard measure of portfolio performance to calculate Sharpe ratios (SR). Assuming quadratic utility, we also measure how much a risk-averse investor is willing to pay for switching from one liquidity measure to another. For each of these economic evaluation metrics, we obtain one ranking of all investigated liquidity measures.

[7] An alternative optimization is to maximize expected utility. We use maximizing expected returns subject to target volatility as it is the most common optimization in the literature. Maximizing expected utility leads to the same ranking of liquidity measures as maximizing expected returns. Results are available from the authors upon request.

[8] Our results are similar when we carry out the out-of-sample asset allocation problem with the covariance matrix predicted by a multivariate GARCH(1,1) model or with the covariance matrix kept constant over time.

Sharpe ratio

The first economic criterion we employ is the Sharpe ratio, or return-to-variability ratio, which measures the risk-adjusted returns from a portfolio or investment strategy and is widely used by investment banks and asset management companies to evaluate investment and trading performance. The ex-post SR is defined as:

$$SR = \frac{\overline{r_s - r_f}}{\sigma_s},$$

where $\overline{r_s - r_f}$ is the average (annualized) excess strategy return over the risk free rate, and σ_s is the (annualized) standard deviation of the investment returns. This measure is commonly used to evaluate performance in the context of mean-variance analysis. However, Marquering and Verbeek (2004) and Han (2006) show that the SR can underestimate the performance of dynamically managed portfolios. This is because the SR is calculated using the average standard deviation of the realized returns, which overestimates the conditional risk (standard deviation) faced by an investor at each point in time. For this reason we use the performance fee as an additional economic criterion to quantify the economic gains from using the liquidity models considered.

Performance fees under quadratic utility

The second economic significance evaluation metric is based on the performance fee. Specifically, we calculate the maximum performance fee a risk-averse investor is willing to pay to switch from the strategy based on liquidity measure A to an alternative strategy that is based on liquidity measure B. This measure is based on mean-variance analysis with quadratic utility (West, Edison, and Cho, 1993; Fleming, Kirby, and Ostdiek, 2001; Rime, Sarno, and Sojli, 2010). Under quadratic utility, at the end of period $t+1$ the investor's utility of wealth can be represented as:

$$U\left(W_{t+1}\right) = W_{t+1} - \frac{\varrho}{2} W_{t+1}^2 = W_t\left(1 + r_{s,t+1}\right) - \frac{\varrho}{2} W_t^2 \left(1 + r_{s,t+1}\right)^2,$$

where W_{t+1} is the investor's wealth at $t+1$; $r_{s,t+1}$ is the gross strategy return; and ϱ determines her risk preference. To quantify the economic value of each model the degree of relative risk aversion (RRA) of the investor is set to $\delta = \frac{\varrho W_t}{1-\varrho W_t}$, and the same amount of wealth is invested every day. Under these conditions, West, Edison, and Cho (1993) show that the average realized utility ($\overline{U}$) can be used to consistently estimate the expected utility generated from a given level of initial wealth. The average utility for an investor with initial wealth $W_0 = 1$ is:

$$\overline{U} = \frac{1}{T}\sum_{t=0}^{T-1}\left(1 + r_{s,t+1} - \frac{\delta}{2(1+\delta)}(1 + r_{s,t+1})^2\right).$$

At any point in time, one set of estimates of the conditional returns is better than a second set if investment decisions based on the first set leads to higher average realized utility, $\overline{U}$. Alternatively, the optimal model requires less wealth to yield a given level of $\overline{U}$ than a suboptimal model. Following Fleming, Kirby, and Ostdiek (2001), we measure the economic value of liquidity by equating the average utilities for selected pairs of portfolios. Suppose, for example, that holding a portfolio constructed using the optimal weights based on liquidity measure A yields the same average utility as holding the optimal portfolio implied by the liquidity measure B that is subject to daily expenses Φ, expressed as a fraction of wealth invested in the portfolio. Since the investor would be indifferent between these two strategies, we interpret Φ as the maximum performance fee she will pay to switch from strategy A to strategy B. In other words, this utility-based criterion measures how much a mean-variance investor is willing to pay for conditioning on a particular liquidity measure for the purpose of forecasting stock returns. The performance fee will depend on the investor's degree of risk aversion. To estimate the fee, we find the value of Φ that satisfies:

$$\sum_{t=0}^{T-1}\left\{1 + r^A_{s,t+1} - \frac{\delta}{2(1+\delta)}\left(1 + r^A_{s,t+1}\right)^2\right\} = \sum_{t=0}^{T-1}\left\{\left(1 + r^B_{s,t+1} - \Phi\right) - \frac{\delta}{2(1+\delta)}\left(1 + r^B_{s,t+1} - \Phi\right)^2\right\},$$

where $r^A_{s,t+1}$ is the strategy return obtained using forecasts based on the liquidity measure A, Φ is the maximum performance fee an investor wants to pay to switch from strategy A to strategy B, and δ is the degree of relative risk aversion (RRA) of the investor.

Transaction costs

In dynamic investment strategies, where the investor rebalances the portfolio every month, transaction costs can play a significant role in determining returns and comparative utility gains. However, traders charge transaction costs according to counter-party types and trade size. Thus, instead of assuming a fixed cost, we compute the break-even transaction cost τ, which is the minimum monthly proportional cost that cancels the utility advantage of a given strategy. A similar measure of transaction costs has been used by Han (2006), Marquering and Verbeek (2004), and King, Sarno, and Sojli (2010). We assume that transaction costs at time t equal a fixed proportion τ of the amount traded in asset k:

$$\tau \sum_{k=1}^{K} A_{k,t} \left| w_{k,t} - w_{k,t-1} \left(\frac{1 + r_{k,t} + r_{f,t-1}}{1 + r_{s,t}} \right) \right|,$$

where $k = 1, \ldots, K$ refers to the risky assets and $A_{k,t} = \frac{costs_{k,t}}{costs_{K,t}}$ is a scaling factor that expresses the break-even transaction costs τ in terms of asset K. The scaling factor takes into account the difference in trading costs between the different assets. To quantify the transaction cost we use the Effective Tick estimates: $costs_{k,t} =$ *Eff. Tick*$_{k,t}$. The choice for Effective Tick is based on Goyenko, Holden, and Trzcinka (2009) who find that this measure is the best proxy for effective spread, which is an estimate of the execution cost actually paid by the investor. Previous articles assume that the transaction costs of all assets in their analysis is the same, i.e. $A_{k,t} = 1$. We cannot make that assumption because we focus on the liquidity differences between assets, which implies that the transaction costs of the different assets are unlikely to be the same.

2.3 Data

We use daily data of common stocks listed on the New York Stock Exchange (NYSE) from 1947-2008. All data are obtained from the Center for Research in Security Prices (CRSP). We use the daily data to construct the monthly variables. Using daily data, instead of high frequency data, enables us to investigate a longer sample period. Following the literature (see e.g. Chordia, Roll, and Subrahmanyam, 2000; Hasbrouck, 2009; Goyenko, Holden, and Trzcinka, 2009), we include only stocks that have sharecode 10 or 11 and do not change ticker symbol, CUSIP, or primary exchange over the sample period. Days with unusually low volume due to holidays are removed. Our final sample includes 16,083,228 stock/day observations. Table 2.A.1 in the Appendix presents the sample characteristics. The average daily price in the sample is $\$31$ and average trading volume $\$10$ million. The average annualized volatility is 24% and turnover is 74%.

The dependent variable in all our regressions is monthly excess returns. All monthly stock returns are adjusted for delisting bias following Shumway (1997).[9] Excess returns $r^e_{i,t}$ are calculated above the 1-month Treasury bill rate from Ibbotson Associates as provided on Kenneth French's website.[10]

2.3.1 Portfolio construction

We use liquidity and excess return series of size portfolios instead of individual stocks in the regressions. The aggregation of individual stocks into portfolios is necessary to reduce the number of assets in the asset allocation. It also deals with issues related to individual stocks that enter and leave the sample, due to delistings and IPOs. Before aggregating the individual stocks into portfolios, we filter the individual observations based on the level of the stock price, the number of daily observations within the month, and the availability of size, liquidity, and return information.[11] Stock i is included in a portfolio in month t if it satisfies the following criteria:

(1) The preceding month-end stock price is between \$5 and \$1,000 $\left(5 < p_{i,t-1,D_{i,t-1}} < 1000\right)$, where $p_{i,t-1,D_{i,t-1}}$ is the stock price of stock i on day $D_{i,t-1}$ in month $t-1$. This rules out returns that are affected by the minimum tick size.

(2) The preceding month-end market capitalization information $(M_{i,t-1})$ is available, which we need for sorting.

(3) $LIQ_{i,t-1}$ is available and is computed using at least 15 daily observations to ensure the quality of the measure.

[9] For all delistings we use the delisting returns available in CRSP. If this return is not available and the delisting code is 500, 520, 551-574, 580, or 584, we follow Shumway (1997) and use a return of -30%.

[10] `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

[11] The filtering criteria are in line with Amihud (2002), Pástor and Stambaugh (2003), Acharya and Pedersen (2005), and Ben-Rephael, Kadan, and Wohl (2010).

(4) After excluding individual monthly observations that do not satisfy conditions (1) to (3), we winsorize each month across all remaining stocks to the top and bottom 1% of the liquidity variables to avoid outliers.

After filtering, the sample consists of 4,348 stocks. We sort these stocks based on previous end-of-month market capitalization in $K = 10$ size portfolios. The portfolio liquidity and return series are simply the cross-sectional averages of the included individual stocks.

Directly sorting on the liquidity measure of interest is not possible because we analyze various liquidity measures, which would lead to different rankings and different portfolio components for each measure. If we construct different portfolios for each individual liquidity measure, we will not be able to disentangle whether performance differences are due to the composition of the portfolio or to the better predictive quality of the liquidity measure. Furthermore, Amihud (2002) finds that the effect between liquidity and expected returns is stronger for small firms than for large firms. By creating portfolios based on size we take this into account in the econometric framework.

2.3.2 Liquidity variables

We consider a variety of monthly liquidity measures, which can be constructed over the entire sample period 1947–2008 and together capture all aspects of liquidity: Roll, Effective Tick, Zeros, High-Low, and Illiquidity Ratio (ILR).[12] The first four measures proxy for the bid-ask spread and the fifth measure is a proxy for price impact. All liquidity variables measure illiquidity, i.e. higher estimates correspond to lower liquidity.

Roll

Roll (1984) shows that trading costs lead to a negative serial correlation in subsequent price changes. In other words, the effective bid-ask spread is inversely related to the covariance between subsequent price changes. The Roll measure is calculated as:

$$\mathit{Roll}_{i,t} = \begin{cases} 2\sqrt{-Cov\left(\Delta p_{i,t,d};\Delta p_{i,t,d-1}\right)}, & \text{if } Cov\left(\Delta p_{i,t,d};\Delta p_{i,t,d-1}\right) < 0, \\ 0, & \text{if } Cov\left(\Delta p_{i,t,d};\Delta p_{i,t,d-1}\right) \geq 0. \end{cases}$$

[12] Some liquidity measures that we leave out are: the measures developed in Chordia, Huh, and Subrahmanyam (2009) because they require analyst data, and the Sadka (2006) measure based on high-frequency data.

where $\Delta p_{i,t,d}$ is the price change for stock i in month t on day d with $d = 1, 2, \ldots, D_{i,t}$ and $D_{i,t}$ the total number of trading days of stock i in month t.

Effective tick

Holden (2009) and Goyenko, Holden, and Trzcinka (2009) jointly develop a liquidity measure based on price clustering, which builds on the findings of Harris (1991) and Chrisie and Schultz (1994). If one assumes that the spread size is the only cause of price clustering, observable price clusters can be used to infer the spread. If prices are exclusively quoted on even eight increments $\left(\$\frac{1}{4}, \$\frac{1}{2}, \$\frac{3}{4}, \$1\right)$ the spread must be $\$\frac{1}{4}$ or larger. However when prices are also quoted on odd eight increments $\left(\$\frac{1}{8}, \$\frac{3}{8}, \$\frac{5}{8}, \$\frac{7}{8}\right)$ the spread must be $\$\frac{1}{8}$. If the minimum tick size is $\$\frac{1}{8}$, there are $J = 4$ possible spreads: $s_1 = \$\frac{1}{8}; s_2 = \$\frac{1}{4}; s_3 = \$\frac{1}{2}; s_4 = \$1$. The observed fraction F_j of odd $\$\frac{1}{8}, \$\frac{1}{4}, \$\frac{1}{2}, \$1$ prices can be used to estimate the probability γ_j of a certain spread s_j. The unconstrained probability $U_{i,t,j}$ of the j^{th} spread s_j for stock i in month t is:

$$U_{i,t,j} = \begin{cases} 2F_{i,t,j} & \text{if } j = 1 \\ 2F_{i,t,j} - F_{i,t,j-1} & \text{if } j = 2, 3, \ldots, J_t - 1 \\ F_{i,t,j} - F_{i,t,j-1} & \text{if } j = J_t, \end{cases}$$

where $F_{i,t,j}$ is the observed fraction of trades on prices corresponding to the j^{th} spread for stock i in month t: $F_{i,t,j} = \frac{N_{i,t,j}}{\sum_{j=1}^{J_t} N_{i,t,j}}$ for $j = 1, 2, \ldots, J_t$. with $N_{i,t,j}$ the number of positive volume days in month t that correspond to the j^{th} spread. The unconstrained probabilities $U_{i,t,j}$ can be below zero or above one, so we add restrictions to make sure the $\gamma_{i,t,j}$'s are real probabilities:

$$\gamma_{i,t,j} = \begin{cases} \min[\max(U_{i,t,j}, 0), 1] & \text{if } j = 1 \\ \min[\max(U_{i,t,j}, 0), 1 - \sum_{m=1}^{j-1} \gamma_{i,t,m}] & \text{if } j = 2, 3, \ldots, J_t. \end{cases}$$

The effective tick measure is the expected spread scaled by the average price over that month.

$$\textit{Eff. Tick}_{i,t} = \frac{\sum_{j=1}^{J_t} \gamma_{i,t,j} s_{i,t,j}}{\overline{p}_{i,t}}.$$

Zeros

Lesmond, Ogden, and Trzcinka (1999) develop a liquidity measure based on the proportion of days with zero returns. In a day with zero return, the value of trading on information does not exceed transaction costs for an investor on that day. A less liquid asset with high transaction costs is less often traded than a more liquid asset, and the less liquid asset has a higher proportion of days with zero returns. Zeros is measured as:

$$\textit{Zeros}_{i,t} = \frac{\sum_{d=1}^{D_{i,t}} I_{\{r_{i,t,d}=0\}}}{D_{i,t}},$$

where $I_{\{r_{i,t,d}=0\}}$ is an indicator function that takes the value 1 if the return of stock i on day d in month t is zero.

High-low

Corwin and Schultz (2012) develop a liquidity measure based on daily high and low prices. It is likely that the highest price on a particular day is against the ask quote while the lowest price is against the bid quote. The high-low ratio therefore reflects both a stock's variance and its bid-ask spread. To disentangle the variance and the spread component, we make use of multiple time intervals since the variance is proportional to the length of the interval, while the spread component is not. Hence, the liquidity measure is written as a function of one-day and two-day high-low ratios:

$$\textit{High-Low}_{i,t} = \frac{2\left(\exp^{\alpha} - 1\right)}{1 + \exp^{\alpha}},$$

where $\alpha = \frac{\sqrt{2\beta}-\sqrt{\beta}}{3-2\sqrt{2}} - \sqrt{\frac{\gamma}{3-2\sqrt{2}}}$, $\beta = \log\left(\frac{H_{i,t}}{L_{i,t}}\right)^2 + \log\left(\frac{H_{i,t-1}}{L_{i,t-1}}\right)^2$, $\gamma = \log\left(\frac{H_{i,t-1,t}}{L_{i,t-1,t}}\right)^2$, $H_{i,t}\ (L_{i,t})$ is the high (low) price on day t for asset i, and $H_{i,t-1,t}\ (L_{i,t-1,t})$ is the high (low) price over the two days $t-1$ and t for asset i.

Illiquidity ratio

The measure developed in Amihud (2002) proxies for the price impact of a trade. Price impact refers to the positive relation between transaction volume and price change. The measure is defined as the ratio between the absolute daily return over dollar volume, averaged over the month:

$$ILR_{i,t} = \frac{1}{D_{i,t}} \sum_{d=1}^{D_{i,t}} \frac{|r_{i,t,d}|}{V_{i,t,d}},$$

where $V_{i,t,d}$ is the dollar volume of traded stocks i (in millions) on day d and $D_{i,t}$ is the number of trading days in month t. To be able to compare the ILR over time, we correct it for inflation and the increased size of financial markets. When adjusting the series we cannot use future information that is not available to a real-time investor. We follow Acharya and Pedersen (2005) and Pástor and Stambaugh (2003) and scale the liquidity measure by a ratio of market capitalizations:

$$ILR_{i,t}^{adj} = ILR_{i,t} \frac{M_{m,t-1}}{M_{m,1}},$$

where $ILR_{i,t}$ is the illiquidity ratio in month t of stock i and $M_{m,t-1}$ is the market capitalization in month $t-1$ and $M_{m,1}$ is the market capitalization in January 1947. In the remainder of this paper we drop the superscript and refer to the adjusted Amihud illiquidity ratio with *ILR*.

2.3.3 Preliminary statistics

Table 2.1 presents the liquidity characteristics for the market and three size portfolios. Panel A shows the liquidity characteristics for the market portfolio. Panels B, C, and D show the liquidity characteristics for the size portfolios. The portfolio with small firms (Panel B) is the least liquid and has the most variability over time. The bottom three panels (E - G) show market liquidity over three subperiods of 20 years. The ILR measure shows that price impact is the lowest in the post-war sub-period (1947-1967) with a value of 2.537 and has the lowest volatility of 0.725. In contrast, Zeros shows that liquidity has increased over time.

Table 2.1 Liquidity statistics for size portfolios

The table presents time series characteristics for monthly liquidity series. The sample period is January 1947 to December 2008. Size portfolios are formed based on previous month market capitalization $(M_{i,t-1})$. The liquidity measures are: ILR (Amihud, 2002), Roll (Roll, 1984), Effective Tick (Holden, 2009; Goyenko et al., 2009), Zeros (Lesmond et al., 1999), and High-Low (Corwin and Schultz, 2012). Panel A shows the liquidity characteristics of the equally weighted market portfolio over the entire sample period. Panel B, C, and D show the liquidity characteristics for size portfolio 1, 5, and 10, respectively, where portfolio 1 consists of the smallest firms and portfolio 10 of the largest firms. Panel E, F, and G show the liquidity characteristics of the equally weighted market portfolio for three subperiods.

	ILR	Eff. tick	Roll	Zeros	High-low
Panel A: Market portfolio					
Average	2.904	0.0077	0.0090	0.151	0.0065
Median	2.641	0.0082	0.0083	0.169	0.0062
Volatility	1.358	0.0031	0.0028	0.061	0.0014
Panel B: Size portfolio 1, small firms					
Average	13.294	0.0147	0.0128	0.227	0.0101
Median	11.868	0.0160	0.0124	0.249	0.0101
Volatility	6.628	0.0055	0.0032	0.082	0.0023
Panel C: Size portfolio 5					
Average	1.562	0.0074	0.0089	0.152	0.0063
Median	1.386	0.0079	0.0083	0.168	0.0059
Volatility	0.977	0.0031	0.0031	0.065	0.0015
Panel D: Size portfolio 10, large firms					
Average	0.118	0.0034	0.0062	0.092	0.0047
Median	0.106	0.0036	0.0053	0.098	0.0044
Volatility	0.100	0.0014	0.0034	0.045	0.0017
Panel E: Market portfolio (1947–1967)					
Average	2.537	0.0089	0.0085	0.183	0.0061
Median	2.388	0.0083	0.0080	0.181	0.0059
Volatility	0.725	0.0018	0.0024	0.029	0.0009
Panel F: Market portfolio (1968–1988)					
Average	3.258	0.0092	0.0086	0.172	0.0067
Median	2.743	0.0090	0.0083	0.171	0.0064
Volatility	1.752	0.0018	0.0018	0.028	0.0010
Panel G: Market portfolio (1989–2008)					
Average	2.916	0.0049	0.0098	0.096	0.0068
Median	2.910	0.0049	0.0086	0.080	0.0061
Volatility	1.296	0.0033	0.0038	0.074	0.0020

Figure 2.1 shows the time series of the liquidity measures. The Roll and High-Low liquidity measures are quite stable and fluctuate around means of 0.01 and 0.006 respectively. Both measures show low liquidity around 1975 (oil crisis), 2000 (burst of internet bubble), and 2007-2008 (global financial crisis). The Effective Tick and Zeros measures exhibit a

decreasing trend. There are two permanent shocks in these series that coincide with the two minimum tick changes: on June 24, 1997, the minimum tick decreases from $\frac{1}{8}$ to $\frac{1}{16}$, and on January 29, 2001, it decreases from $\frac{1}{16}$ to 0.01. The last measure, ILR, shows periods of illiquidity in 1970, 1975, in the beginning of the 90's, in 2000, and in 2008. The figure shows that it is important to allow the relation between liquidity and conditional expected returns to vary over time by using rolling estimation windows.

Figure 2.1 Liquidity measures

The figure shows the liquidity measures for the market over time. The market series is computed as the cross-sectional equally-weighted average of individual stock liquidity measures. The sample period is January 1947 to December 2008. The low-frequency liquidity measures are: ILR (Amihud, 2002), Roll (Roll, 1984), Effective Tick (Holden, 2009; Goyenko et al., 2009), and Zeros (Lesmond et al., 1999), and High-Low (Corwin and Schultz, 2012).

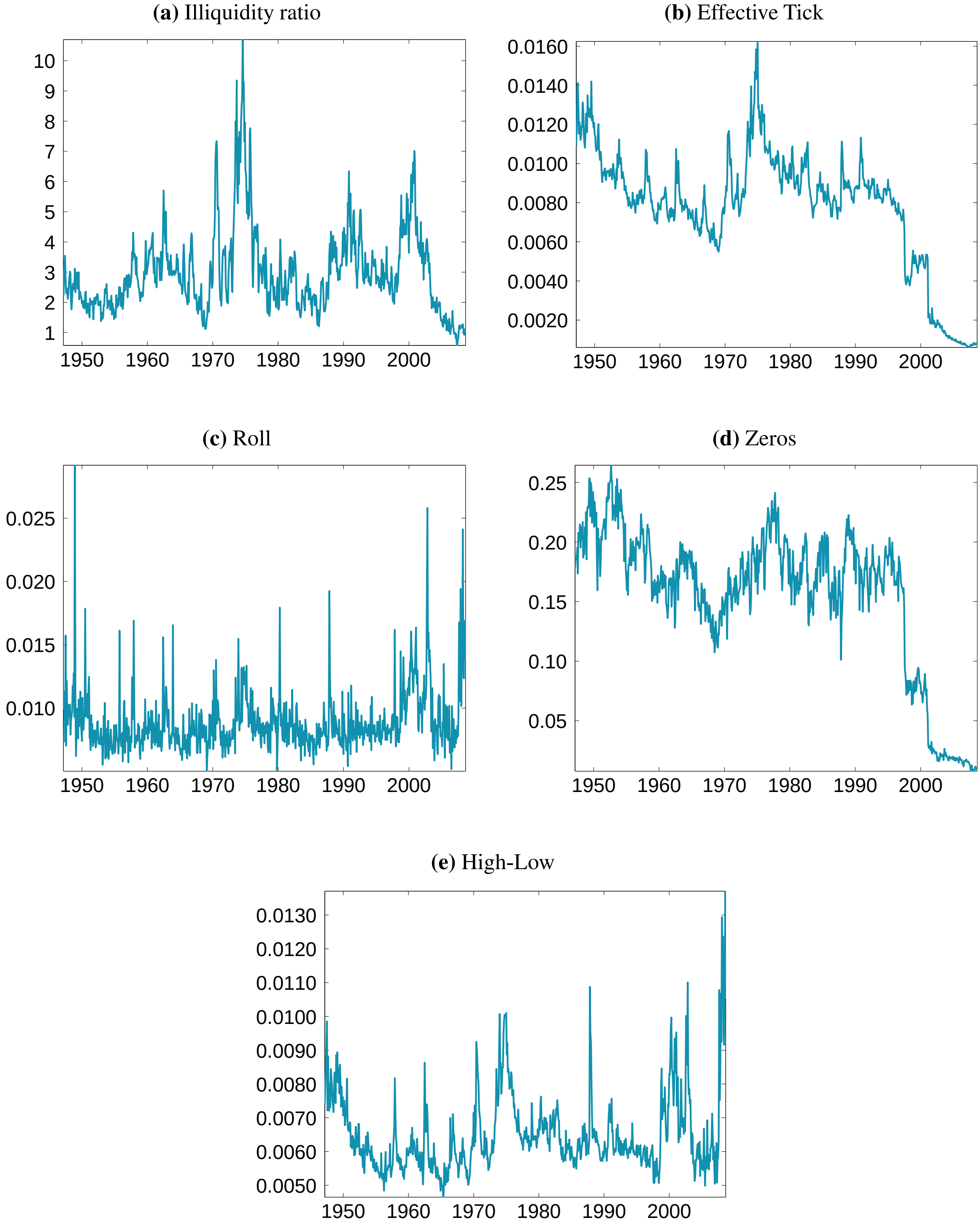


Table 2.2 shows the descriptive statistics for the market portfolio excess returns and three size portfolios, over the entire sample and three subsamples: 1947-1967, 1968-1988, and 1989-2008. In line with Fama and French (1992), we find that small firms have both more volatile and higher average returns compared to the returns for large firms. The different subsamples show that returns vary considerably over time. The preliminary statistics provide initial evidence of a possible link between liquidity and returns.

Table 2.2 Excess return statistics for size portfolios

The table presents descriptive characteristics for monthly excess returns. All reported values are annualized. The sample period is January 1947 to December 2008 and is divided into three subperiods. Size portfolios are formed based on previous month market capitalization $(M_{i,t-1})$. Panel A shows the characteristics of the equally weighted market portfolio. Panel B, C, and D show the characteristics for size portfolio 1, 5, and 10, respectively, where portfolio 1 consists of the smallest firms and portfolio 10 of the largest firms.

	Entire sample	1947-1967	1968-1988	1989-2008
Panel A: Market portfolio				
Average	7.50%	12.28%	4.98%	5.30%
Median	12.58%	17.17%	5.23%	10.03%
Volatility	17.06%	14.22%	20.45%	15.82%
Panel B: Size portfolio 1, small firms				
Average	8.42%	13.72%	9.24%	2.32%
Median	12.29%	16.10%	12.03%	7.22%
Volatility	20.47%	17.86%	24.50%	18.21%
Panel C: Size portfolio 5				
Average	7.75%	12.28%	5.14%	5.90%
Median	11.56%	17.86%	5.85%	10.14%
Volatility	18.12%	14.98%	21.56%	17.18%
Panel D: Size portfolio 10, large firms				
Average	5.19%	10.24%	1.51%	3.96%
Median	10.17%	13.85%	4.01%	9.45%
Volatility	14.35%	12.18%	16.65%	13.77%

2.4 Results

We now turn to the results of our main analysis. Here, all liquidity series are modeled using an autoregressive model of order two. The optimal number of lags is based on the Akaike Information Criterion and the Bayesian information Criterion. In this section, all switching fees are based on a relative risk aversion coefficient $\delta = 5$, and the ex-ante target volatility is set to $\sigma_s^* = 10\%$. In Section 2.6 we examine the robustness of the results to these settings.

2.4.1 Main results

Table 2.3 shows the performance of the liquidity timing strategies. We present the following performance measures: the Sharpe ratio (SR, column 1), the relative performance expressed in switching fees (columns 2-5), transaction costs (columns 6-7), and the excess return and volatility (columns 8-9). Each row corresponds to the characteristics of a strategy that conditions on a particular liquidity measure. The results in Panel A indicate that it is possible to use liquidity timing to earn positive returns, and that the Zeros measure performs best. The SR of the Zeros strategy is 0.38 and is higher than the SR of ILR (0.13), Roll (-0.04), Effective Tick (-0.07), and High-Low (-0.16).[13]

The Zeros strategy performs best as indicated by the positive switching fees of all variables towards Zeros. A risk-averse investor would pay 287.3 basis points per year to switch from the ILR strategy to a strategy that conditions on Zeros. The High-Low strategy has the worst performance because a risk-averse investor does not want to pay a positive fee to switch to the High-Low strategy.[14] The Zeros break-even transaction costs (τ_1) are 4.2 basis points, if we assume that transaction costs are the same for all risky assets. When we incorporate the cost differences and express the break-even transaction costs in terms of the most liquid asset, we find $\tau_A = 2.0$ basis points. The next two columns show the excess returns and their volatilities. The excess return of Zeros is the highest, 4.55%.

Panel B shows the strategy characteristics when the return predictions are based on the average return prediction using three different rolling windows (5, 10, and 20 years). The Zeros strategy has both the highest excess return (6.04%) and SR (0.51). The positive switching fees for the Zeros strategy show that a risk-averse investor always wants to pay a positive fee to switch to condition on the Zeros measure. Compared to Panel A, the excess returns in Panel B are higher for four of the five measures and the return volatilities remain similar. Averaging the return predictions of different rolling windows seems to deliver better performance, which could be related to more accurate return forecasts, in line with Pesaran and Pick (2011).

[13]The differences in SR are not only economically but also statistically significant. According to the SR test of Ledoit and Wolf (2008), the SR of the Zeros strategy is significantly higher than the other SRs, except for the SR of the ILR strategy.

[14]The break-even transaction costs are not computed for the Effective Tick, Roll, and High-Low strategy because their excess returns are negative.

Table 2.3 Performance of dynamic asset allocation strategies
The table presents dynamic strategy results for the different liquidity measures. The dynamic asset allocation strategies are based on a mean-variance framework where the investor maximizes the conditional expected return subject to a target conditional volatility ($\sigma_s^* = 10\%$). The return predictions of the risky assets in Panel A are based on a 10-year rolling window and in Panel B are based on the average prediction of a 5, 10, and 20 year rolling window. The weights of the risky assets are not restricted. The sample period is January 1947 to December 2008, and the strategies start trading in January 1967. The liquidity measures are: ILR (Amihud, 2002), Roll (Roll, 1984), Effective Tick (Holden, 2009; Goyenko et al., 2009), Zeros (Lesmond et al., 1999), and High-Low (Corwin and Schultz, 2012). All numbers are annual, except for the break-even costs τ_1 and τ_A that are reported in basis points per trade. The switching fee is the maximum performance fee a risk-averse investor is willing to pay to switch from one strategy to another. It is expressed in annual basis points and is computed based on a relative risk aversion parameter of 5. τ_1 shows the break-even costs under the assumption that all assets have the same transaction costs. τ_A shows the break-even costs in terms of large firms, whereby the liquidity differences between small and large firms are taken into account. If the excess return of a strategy is negative we do not compute break-even transaction costs and report the symbol "-".

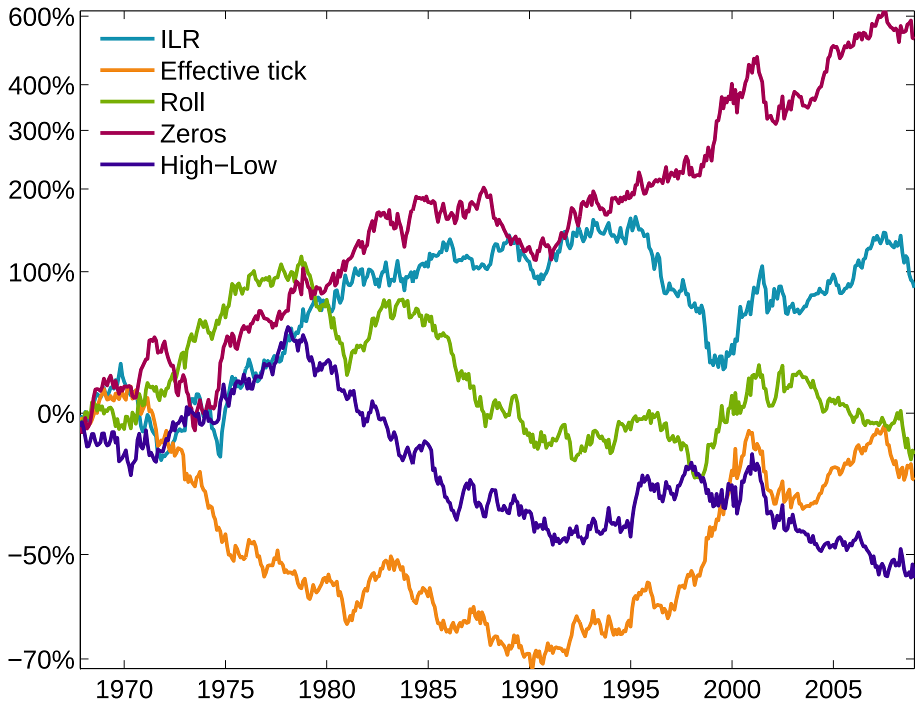
		Switching fee							
	SR	ILR	Eff. tick	Roll	Zeros	τ_1	τ_A	Excess return	Volatility
Panel A: 10 year window									
ILR	0.13					1.4	0.6	1.50%	11.88%
Eff. tick	−0.07	−206.2				–	–	−0.79%	11.41%
Roll	−0.04	−192.2	13.8			–	–	−0.49%	11.77%
Zeros	0.38	287.3	494.1	480.0		4.2	2.0	4.55%	12.05%
High-low	−0.16	−349.7	−143.7	−157.5	−637.5	–	–	−1.94%	11.99%
Panel B: Combination of windows									
ILR	0.27					2.8	1.3	3.14%	11.54%
Eff. tick	−0.13	−477.7				–	–	−1.54%	11.81%
Roll	0.06	−245.6	232.5			0.8	0.4	0.72%	11.69%
Zeros	0.51	264.8	742.5	510.0		5.5	2.6	6.04%	11.79%
High-low	−0.09	−455.6	22.5	−210.0	−720.0	–	–	−1.15%	12.15%

Figure 2.2 shows the cumulative returns of all five strategies based on the 10 year rolling window predictions. The best performing strategies are based on the ILR and the Zeros measures. The outperformance of the Zeros strategy is not generated during a particular period, since its returns steadily increase over the entire sample period. The ILR strategy performs very well and tracks the Zeros strategy until 1995, but it decreases substantially between 1995 and 1998 and never recovers. The decline in performance of the ILR strategy is in line with Ben-Rephael, Kadan, and Wohl (2010), who argue that the profitability of trading strategies based on volume related liquidity proxies declined over the past four decades. Both the Roll and High-Low strategies show good performance until 1980, but they become loss-making afterwards. Finally, the Effective Tick strategy loses money until 1990, it sharply increases until 2001, and in the final years of the sample its performance is flat.

A liquidity measure that is closely related to Zeros is the LM1 measure of Liu (2006). Both measures proxy for liquidity using the number of days where there is no information, however LM1 uses zero volume days while Zeros is based on zero return days. LM1 is computed as the standardized turnover-adjusted number of zero daily trading volumes per month: $LM1 = [NoZV + \frac{1/turnover}{Deflator}]\frac{21}{NoTD}$, with $NoZV$ the number of zero daily volumes, $turnover$ the sum of daily ratios of number of traded shares to the number of shares outstanding, $Deflator$ a correction term, and $NoTD$ the number of trading days. The results in Table 2.A.2 in the Appendix show that the ranking of strategies using LM1 is the same as using Zeros. Given the similarity between the two measures, we continue using the Zeros measure as the main object of interest.

Figure 2.2 Cumulative returns of dynamic asset allocation strategies

The figure shows the cumulative log returns of the dynamic asset allocation strategies that use liquidity information to predict excess returns. The low-frequency liquidity measures are: ILR (Amihud, 2002), Roll (Roll, 1984), Effective Tick (Holden, 2009; Goyenko et al., 2009), Zeros (Lesmond et al., 1999), and High-Low (Corwin and Schultz, 2012). The dynamic asset allocation strategies are based on a mean-variance framework where the investor maximizes the conditional expected return subject to a target conditional volatility. The return predictions of the risky assets are based on a 10-year rolling window.



2.4.2 Control variables

This section deals with the possibility that an omitted variable, which is correlated with the liquidity measures, is driving the results. Welch and Goyal (2008) examine the predictive ability of several market return predictors suggested in the existing literature.[15] We use only the Welch and Goyal (2008) predictors that are publicly available for the entire sample period (1947-2008).[16] All variables are constructed for the market and not for the individual size portfolios.

Adding the control variables to Equation (2.2), with $p = 2$ yields:

$$E_t\left[r_{k,t+1} - r_{f,t}\right] = \beta_{0,t} + \beta_{1,t} LIQ_{k,t} + \beta_{2,t} LIQ_{k,t-1} + \sum_{n=1}^{N} \gamma_{n,t} f_{n,t}, \tag{2.4}$$

where $LIQ_{k,t}$ is the liquidity of asset k at time t and $f_{n,t}$ are the $n = 1, 2, \ldots, N$ control variables: dividend yield, earnings price ratio, dividend payout ratio, stock variance, book-to-market ratio, net equity expansion, term-spread, default yield spread, default return spread, and inflation. We estimate the β- and γ-parameters in the same way as in the main analysis. Some of the control variables, especially the dividend price ratio and dividend yield, are highly correlated. Hence, we run separate regressions for each control variable and one regression where we include all control variables, excluding the dividend price ratio to avoid singularity issues.[17]

We test whether any of the control variables yield better predictions compared to the model with also a liquidity term. In this set-up we first estimate a model where expected excess returns are only driven by a constant and one of the control variables. Second, we estimate a model where expected excess returns are driven by a constant, one of the control variables, and a liquidity variable. We then check if the added liquidity variable increases performance, compared to the first case.

The results are presented in Table 2.4. Columns (1) - (3) show the SR and break-even transaction costs if the model consists of a constant and one control variables. Columns (4)

[15] They investigate the dividend price ratio, dividend yield, earnings price ratio, dividend payout ratio, stock variance, cross-sectional premium, book-to-market ratio, net equity expansion, percent equity issuing, term-spread, default yield spread, default return spread, inflation, and investment to capital ratio.

[16] We exclude percent equity issuing and investment to capital ratio because they are not publicly available. All available data are from Amit Goyal's website `http://www.hec.unil.ch/agoyal/`.

[17] We include the dividend yield because it shows the highest SR on an individual basis (see Table 2.4). Furthermore, replacing the dividend yield by the dividend price ratio gives similar results.

- (8) show the SR when liquidity is added to predict returns. When considering the first 10 control variables, the highest SR in column (1) of Panel A is 0.21, which is lower than the highest SR of a strategy conditioning only on liquidity (0.38 from the Zeros strategy in Table 3). Comparing column (7) with column (1), shows that adding the Zeros measure always increases the SR. The only exception is when we add Zeros to a strategy consisting of all control variables, in which case the SR only improves in Panel B. The Zeros measure is followed by the ILR measure, which increases the SR in 6 out of 12 cases. In Panel B the best control variable also underperforms the best liquidity strategy, the SR is 0.28 versus 0.51. The last rows of both Panel A and B show the performance of the net equity expansion variable (ntis). In contrast to the other 10 controls, this variable performs really well with an SR of 0.42 in both panels. According to Baker and Stein (2004) this variable is related to liquidity. Hence, this supports our idea that liquidity has predictive ability for returns.

Table 2.4 Performance of dynamic asset allocation strategies for single control variables
The table presents dynamic strategy results for the different control variables. Columns (1) - (3) show results of strategies that condition on a constant and only one of the 11 control variables: dividend price ratio (d/p), dividend yield (d/y), earnings price ratio (e/p), dividend payout ratio (d/e), stock variance (svar), book-to-market ratio (b/m), term-spread (tms), default yield spread (dfy), default return spread (dfr), inflation (infl), and net equity expansion (ntis). The last strategy (all) contains all control variables, except the dividend price ratio. Columns (4) - (8) show Sharpe ratios of strategies that condition both on a constant, one control variable, and one liquidity variable. The dynamic asset allocation strategies are based on a mean-variance framework where the investor maximizes the conditional expected return subject to a target conditional volatility ($\sigma_s^* = 10\%$). The return predictions of the risky assets in Panel A are based on a 10-year rolling window and in Panel B are based on the average prediction of a 5, 10, and 20 year rolling window. The weights of the risky assets are not restricted. The sample period is January 1947 to December 2008, and the strategies start trading in January 1967. All numbers are annual, except for the break-even costs τ_1 and τ_A that are reported in basis points per trade. τ_1 shows the break-even costs under the assumption that all assets have the same transaction costs. τ_A shows the break-even costs in terms of large firms, whereby the liquidity differences between small and large firms are taken into account.

	Without liquidity			With liquidity				
	SR	τ_1	τ_A	ILR	Eff. tick	Roll	Zeros	High-low
Panel A: 10 year window								
d/p	0.13	4.7	2.2	0.21	−0.12	−0.10	0.19	−0.22
d/y	0.19	7.1	3.3	0.23	−0.11	−0.06	0.21	−0.14
e/p	0.08	3.0	1.4	0.15	0.06	−0.13	0.26	−0.23
d/e	0.14	5.1	2.3	0.16	−0.05	−0.06	0.42	−0.20
svar	0.04	1.3	0.6	0.16	0.06	−0.09	0.22	−0.30
b/m	0.02	0.9	0.4	0.08	0.04	−0.13	0.25	−0.14
tms	0.21	6.9	3.2	0.11	−0.12	−0.08	0.36	−0.26
dfy	0.21	7.1	3.3	0.18	0.04	−0.00	0.26	−0.11
dfr	0.17	2.9	1.3	0.17	−0.05	−0.00	0.39	−0.12
infl	0.11	2.3	1.1	0.06	−0.03	−0.04	0.37	−0.17
ntis	0.42	14.5	6.7	0.22	−0.01	0.14	0.47	−0.02
all	0.30	4.6	2.2	0.27	−0.05	0.02	0.21	−0.25
Panel B: Combination of windows								
d/p	0.19	6.7	3.1	0.20	−0.13	−0.08	0.42	−0.18
d/y	0.28	9.5	4.4	0.28	−0.05	−0.02	0.43	−0.10
e/p	0.15	5.0	2.3	0.29	0.07	−0.07	0.44	−0.21
d/e	0.09	3.6	1.7	0.33	−0.05	0.06	0.54	−0.16
svar	0.06	1.9	0.9	0.27	−0.04	0.03	0.34	−0.24
b/m	0.07	2.7	1.3	0.16	−0.04	−0.12	0.47	−0.12
tms	0.18	5.8	2.7	0.29	−0.07	0.03	0.45	−0.20
dfy	0.26	8.3	3.8	0.22	−0.03	0.06	0.44	−0.07
dfr	0.15	2.5	1.2	0.29	−0.11	0.07	0.46	−0.07
infl	0.14	3.2	1.5	0.24	−0.11	0.06	0.47	−0.10
ntis	0.42	13.6	6.2	0.40	−0.03	0.24	0.62	−0.02
all	0.25	3.6	1.7	0.21	0.10	−0.01	0.30	−0.26

To summarize, the addition of alternative return predictors shows that the results are not driven by an omitted variable bias. When we condition only on one individual control variable, we do not find a strategy that gets close to the best liquidity strategy in terms of SR.

Furthermore, the Zeros liquidity measure always increases the performance of a strategy that conditions on one control variable.

2.4.3 Cross-sectional predictors

We extend the analysis in the previous section with cross-sectional predictors, i.e. the three Fama-French factors and the Carhart momentum factor. Fama and French (1993) and Carhart (1997) show that these factors can explain cross-sectional return differences. The data are from Wharton Research Database Services (WRDS).

In line with Equation (2.4), we get:

$$E_t\left[r_{k,t+1} - r_{f,t}\right] = \beta_{0,t} + \beta_{1,t} LIQ_{k,t} + \beta_{2,t} LIQ_{k,t-1} + \sum_{n=1}^{M} \kappa_{m,t} h_{m,t},$$

where $LIQ_{k,t}$ is the liquidity of asset k at time t and $h_{m,t}$ are the $m = 1, 2, \ldots, M$ control variables: excess market return, Small-minus-Big, High-minus-Low, and momentum. Table 2.5 shows the results of the strategies when conditioning on only a constant and one control variable and when conditioning on a constant, a control variable, and a liquidity variable. The SR of the excess market return strategy in column (1) of Panel A is 0.41, which is higher than the SR of the Zeros strategy. When we include all four control variables in one strategy, the SR is 0.47. All Sharpe ratios increase when we add the Zeros liquidity measure in Column (7). In other words, the Zeros liquidity measure contains relevant information that increases the quality of the return predictions. The cross-sectional predictors in Panel B do not outperform the Zeros strategy in terms of SR. In addition the SR of the “all strategy” increases when we add the Zeros liquidity measure.

Table 2.5 Performance of dynamic asset allocation strategies for single cross-sectional control variables
The table presents dynamic strategy results for the different control variables. Columns (1) - (3) show results of strategies that condition on a constant and only one of the 4 cross-sectional control variables: excess market return (Mkt-Rf), Small-minus-Big (SMB), High-minus-Low (HML), and momentum (Mom). The last strategy (all) contains all control variables. Columns (4) - (8) show Sharpe ratios of strategies that condition both on a constant, one control variable, and one liquidity variable. The dynamic asset allocation strategies are based on a mean-variance framework where the investor maximizes the conditional expected return subject to a target conditional volatility ($\sigma_s^* = 10\%$). The return predictions of the risky assets in Panel A are based on a 10-year rolling window and in Panel B are based on the average prediction of a 5, 10, and 20 year rolling window. The weights of the risky assets are not restricted. The sample period is January 1947 to December 2008, and the strategies start trading in January 1967. All numbers are annual, except for the break-even costs τ_1 and τ_A that are reported in basis points per trade. τ_1 shows the break-even costs under the assumption that all assets have the same transaction costs. τ_A shows the break-even costs in terms of large firms, whereby the liquidity differences between small and large firms are taken into account. If the excess return of a strategy is negative we do not compute break-even transaction costs and report the symbol "-".

	Without liquidity			With liquidity				
	SR	τ_1	τ_A	ILR	Eff. tick	Roll	Zeros	High-low
Panel A: 10 year window								
Mkt-Rf	0.41	7.4	3.4	0.32	0.20	0.04	0.57	−0.13
SMB	0.21	4.2	1.9	0.13	−0.03	−0.02	0.37	−0.22
HML	0.08	1.8	0.8	0.18	−0.12	−0.10	0.37	−0.20
Mom	0.34	6.9	3.2	0.11	−0.05	0.03	0.47	−0.14
all	0.47	6.0	2.8	0.31	0.20	0.11	0.59	−0.15
Panel B: Combination of windows								
Mkt-Rf	0.38	7.1	3.2	0.36	0.12	0.16	0.72	−0.01
SMB	0.15	3.2	1.4	0.24	−0.07	0.09	0.50	−0.13
HML	−0.03	–	–	0.36	−0.16	0.01	0.48	−0.10
Mom	0.28	5.8	2.7	0.25	−0.07	0.13	0.55	−0.08
all	0.37	5.0	2.3	0.34	0.14	0.22	0.68	−0.03

2.4.4 Risk adjusted returns

Until now, we have compared the different liquidity measures based on excess returns. It is possible that some of the strategies load more on risk than others, and therefore achieve higher returns. In this section we adjust the returns of the strategies for their exposure to the three Fama-French factors, the Carhart Momentum factor, and the Pastor and Stambaugh liquidity factor. Then, we compare the liquidity strategies based on alpha.

The methodology to compute alphas is similar to Brennan, Chordia, and Subrahmanyam (1998), Chordia, Subrahmanyam, and Anshuman (2001b), and Ben-Rephael, Kadan, and Wohl (2010). In the first step we estimate the sensitivity (β^i) of the strategy returns to each risk factor:

$$r_{s,t} - r_{f,t} = \alpha_{s,t} + \sum_i \beta_{s,t}^i F_t^i + \varepsilon_{s,t},$$

where $r_{s,t} - r_{f,t}$ is the excess return at time t of strategy s, $\alpha_{s,t}$ is the risk adjusted return, $\beta^i_{s,t}$ is the sensitivity to risk factor i, and F^i_t are the risk factors. The factor loadings are estimated over the preceding 60 months: $t-60$ to $t-1$. Next, we calculate the alpha as the excess return of the strategy s in month t minus the just estimated loadings on the risk factors multiplied with the realized returns in month t on the risk factors:

$$\alpha_{s,t} = r_{s,t} - r_{f,t} - \widehat{\beta}^{MKT}_{s,t} r_{MKT,t} - \widehat{\beta}^{SMB}_{s,t} r_{SMB,t} - \widehat{\beta}^{HML}_{s,t} r_{HML,t} - \widehat{\beta}^{UMD}_{s,t} r_{UMD,t} - \widehat{\beta}^{LIQ}_{s,t} r_{LIQ,t}. \quad (2.5)$$

Table 2.6 shows the risk-adjusted results of all liquidity strategies. In both Panel A and B the alpha of the Zeros strategy is the highest, $\alpha = 5.79\%$ and $\alpha = 7.01\%$ respectively. The Zeros strategy is followed by the ILR strategy with an alpha of 1.64% in Panel A and 3.36% in Panel B. Also both these alphas are significantly different from zero. The Effective Tick, Roll, and High-Low strategy have alphas that are either zero or negative.

Table 2.6 Risk-adjusted results

The table presents the risk-adjusted results for the different liquidity strategies. The dynamic asset allocation strategies are based on a mean-variance framework where the investor maximizes the conditional expected return subject to a target conditional volatility ($\sigma^*_s = 10\%$). The return predictions of the risky assets in Panel A are based on a 10-year rolling window and in Panel B are based on the average prediction of a 5, 10, and 20 year rolling window. The weights of the risky assets are not restricted. The sample period is January 1947 to December 2008, and the strategies start trading in January 1967. The liquidity measures are: ILR (Amihud, 2002), Roll (Roll, 1984), Effective Tick (Holden, 2009; Goyenko et al., 2009), Zeros (Lesmond et al., 1999), and High-Low (Corwin and Schultz, 2012). We calculate the alpha as the excess return of the strategy minus the estimated loadings on the risk factors multiplied with the realized returns on the risk factors, see Equation 2.5 on page 35. The sensitivity of the strategy returns to each risk factor is estimated over the preceding 60 months. The risk factors that we take into account are: the excess market return, the SMB factor, the HML factor, the Carhart (1997) momentum factor, and the Pástor and Stambaugh (2003) traded liquidity factor. All numbers are annualized.

	Alpha	Volatility	t-stat
Panel A: 10 year window			
ILR	1.64%	12.94%	2.65
Eff. tick	−1.14%	11.80%	−2.01
Roll	−0.75%	12.47%	−1.25
Zeros	5.79%	12.72%	9.50
High-low	−3.58%	12.62%	−5.92
Panel B: Combination of windows			
ILR	3.36%	12.48%	5.62
Eff. tick	−2.12%	12.40%	−3.56
Roll	0.70%	12.43%	1.17
Zeros	7.01%	12.63%	11.59
High-low	−2.98%	12.72%	−4.90

The alphas of the Zeros and ILR strategies are higher than the excess returns in Table 2.3, which implies that these strategies do not load on these risk factors. In the end, the ranking of the strategies remains the same: Zeros shows the best performance, followed by the ILR strategy.

2.5 Why Zeros performs best?

In all previous analyses we find that the Zeros strategy outperforms other liquidity strategies. In this section we examine why the liquidity timing strategy based on the Zeros measure outperforms the strategies using other liquidity measures.

2.5.1 Performance conditional on market returns

To get a better understanding of the differences in performance between the liquidity strategies, we investigate how the strategies perform in extreme market conditions. We sort the market returns of the past 50 years and condition on particular quantiles of their empirical distribution. Table 2.7 shows the performance of the liquidity strategies conditional on the worst or best x% market returns. In both Panel A and B, only the Zeros strategy achieves positive returns in all cases. This implies that even if the market is decreasing, the Zeros strategy goes long and short in the right assets and makes a profit. In contrast, the ILR strategy shows negative performance when the market is decreasing but shows larger profits when the market is going up. Concluding, the Zeros measure outperforms the other liquidity strategies because it is the only strategy that achieves positive performance in both bull and bear markets.

Table 2.7 Performance of liquidity strategy conditional on market returns
The table presents the monthly return performance of the dynamic asset allocation strategies conditional on a particular quantile of the market returns empirical distribution. We obtain the market returns empirical distribution by sorting the market returns of the past 50 years. We investigate the performance of the liquidity strategies for the months where the market returns belong to the top and bottom quantiles of their empirical distribution. All presented numbers are monthly.

	Bottom 1%	Bottom 5%	Bottom 10%	Top 10%	Top 5%	Top 1%
Panel A: 10 year window						
ILR	−3.18%	−2.22%	−0.94%	0.83%	2.20%	3.27%
Eff. tick	−0.54%	−1.32%	−0.42%	0.36%	0.35%	0.68%
Roll	−3.09%	−2.09%	−0.08%	0.40%	0.77%	0.49%
Zeros	1.04%	0.26%	0.70%	0.54%	1.29%	2.30%
High-low	−1.97%	−0.29%	0.30%	0.58%	0.78%	0.57%
Panel B: Combination of windows						
ILR	−1.35%	−0.79%	−0.42%	1.00%	1.64%	2.14%
Eff. tick	−0.02%	−1.59%	−0.72%	0.41%	0.20%	−0.13%
Roll	−2.67%	−1.50%	0.18%	0.21%	0.63%	0.46%
Zeros	2.50%	1.22%	1.11%	0.56%	1.14%	1.76%
High-low	−1.43%	−0.12%	0.42%	0.47%	0.96%	1.45%

2.5.2 Quality of predicted returns

Jagannathan and Ma (2003) show that restricting the weights in an asset allocation problem to be nonnegative, reduces the estimation error in the return prediction parameters. Hence, nonnegativity restrictions should improve the quality of the return predictions, which will increase the performance of the optimal portfolios. The rationale behind improved prediction quality due to nonnegativity constraints is that it has similar effects as shrinking the return predictions. However, if the quality of the predictions is already good and cannot simply be improved by shrinkage, the strategy performance will deteriorate when restrictions are imposed. Thus, by imposing the weights in the asset allocation to be nonnegative, we expect to find that shrinkage effects improve strategy performance based on return predictions of low quality and deteriorate strategy performance using predictions of high quality.

The results in Table 2.8 show that imposing nonnegativity constraints on the asset weights lead to higher excess returns, lower return volatility, and higher SRs for all strategies, except the Zeros strategy. This implies that the quality of the underlying return predictions is improved for all strategies, except the Zeros strategy. Possibly, the quality of Zeros' return predictions was already high and shrinkage lowers the informational quality. Combining these results with the findings in the previous section, we can say that the positions of the

Zeros strategy are often right and the strategy is constrained when introducing short selling restrictions. Although the Zeros strategy no longer shows the best performance when weights are restricted to be positive, it is actually a sign of its ability to long and short the right assets in the unrestricted case.

Instead of restricting the weights of individual assets, DeMiguel, Garlappi, Nogales, and Uppal (2009a) propose to constrain the norm of the asset-weight vector. This generalization nests as special case the approach of Jagannathan and Ma (2003) but at the same time allows for more flexibility. We find qualitatively similar results when we restrict the weights using norm restrictions instead of short-sale restrictions.[18]

Table 2.8 Performance of dynamic asset allocation strategies with nonnegative weights

The table presents dynamic strategy results for the different liquidity measures. The dynamic asset allocation strategies are based on a mean-variance framework where the investor maximizes the conditional expected return subject to a target conditional volatility ($\sigma_s^* = 10\%$). The return predictions of the risky assets in Panel A are based on a 10-year rolling window and in Panel B are based on the average prediction of a 5, 10, and 20 year rolling window. The weights of the risky assets are restricted to be between $0 < w < 1$. The sample period is January 1947 to December 2008, and the strategies start trading in January 1967. The liquidity measures are: ILR (Amihud, 2002), Roll (Roll, 1984), Effective Tick (Holden, 2009; Goyenko et al., 2009), Zeros (Lesmond et al., 1999), and High-Low (Corwin and Schultz, 2012). All numbers are annual, except for the break-even costs τ_1 and τ_A that are reported in basis points per trade. The switching fee is the maximum performance fee a risk-averse investor is willing to pay to switch from one strategy to another. It is expressed in annual basis points and is computed based on a relative risk aversion parameter of 5. τ_1 shows the break-even costs under the assumption that all assets have the same transaction costs. τ_A shows the break-even costs in terms of large firms, whereby the liquidity differences between small and large firms are taken into account.

		Switching fee							
	SR	ILR	Eff. tick	Roll	Zeros	τ_1	τ_A	Excess return	Volatility
Panel A: 10 year window									
ILR	0.35					42.2	16.9	3.89%	11.03%
Eff. tick	0.24	−109.7				30.0	12.2	2.53%	10.54%
Roll	0.26	−92.8	17.1			29.5	12.1	2.80%	10.75%
Zeros	0.26	−91.9	18.3	1.2		33.4	14.0	2.78%	10.69%
High-low	0.27	−84.4	25.5	7.5	7.0	31.1	13.8	2.83%	10.63%
Panel B: Combination of windows									
ILR	0.34					40.7	15.9	3.74%	10.87%
Eff. tick	0.24	−97.5				30.1	12.3	2.57%	10.49%
Roll	0.28	−69.4	27.7			32.0	13.0	2.95%	10.72%
Zeros	0.29	−50.6	47.3	18.8		35.6	14.6	3.14%	10.69%
High-low	0.29	−57.2	40.1	11.2	−7.5	33.1	14.1	3.06%	10.69%

[18] We do not present these results to conserve space, but they are available from the authors upon request.

2.6 Robustness

2.6.1 Benchmarks

In this section we show whether the liquidity timing strategies are related to other timing strategies, which we refer to as benchmarks. The first group of benchmarks predicts returns by conditioning on past return information: (i) the historical average return (Prevailing mean strategy) and (ii) the historical average return and a lagged return term (Lagged return strategy). The second group of benchmarks consists of an equally weighted, a volatility timing, and a minimum variance strategy.

Some of the benchmarks that we use are related to existing literature. The prevailing mean model is used in Welch and Goyal (2008) and Campbell and Thompson (2008). A simple extension to the prevailing mean model is the addition of a lagged return term. If monthly returns are correlated over time this term would improve the quality of return predictions. Both benchmarks closely follow the methodology of the liquidity strategies, only the expression in Equation (2.2) does not contain liquidity variables.

The benchmarks in the second group obtain weights using a different optimization problem than in Equation (2.3). The equally weighted strategy simply gives all risky assets the same weight. Hence, this strategy is long-only and closely resembles the market portfolio. DeMiguel, Garlappi, and Uppal (2009b) show that sample-based mean-variance models have difficulties outperforming such a naive 1/N portfolio.

The second is a volatility timing strategy, similar to Fleming, Kirby, and Ostdiek (2001). This strategy minimizes the portfolio variance, subject to an ex-ante target portfolio return. To eliminate possible predictive power from the return predictions, we set these predictions equal to a constant, i.e. the sample average return of the risky assets. This ensures that the variation in investment weights of this strategy is fully determined by changes in the conditional covariance matrix. Note that we introduce a look-ahead bias by using the average return over the entire sample, which should improve the performance of this benchmark. The last benchmark is a minimum variance strategy. Similar to the volatility timing strategy, the asset weights depend only on the conditional covariance matrix.

Table 2.A.3 in the Appendix shows that the SR of the benchmarks is lower than the SR of the best performing liquidity strategies in Table 2.3. The prevailing mean model has a

low SR, which is in contrast to the findings of Welch and Goyal (2008). This difference can possibly be explained by the differences in data. We use a shorter data sample and we predict the returns of size portfolios, not of the market portfolio. Furthermore, we make use of a rolling window approach whereas they use an expanding window. The addition of a lagged return term leads to a higher SR in Panels A and B.

Panel C shows the performance of the three strategies that obtain investment weights in a different way. The good performance of the equally weighted strategy is in line with the findings in DeMiguel, Garlappi, and Uppal (2009b). Its SR is 0.28 and its turnover is lower than that of the other strategies, as reflected by the high break-even transaction costs. The results of the volatility timing strategy are worse than the liquidity timing strategies. The last row indicates that minimizing volatility also gives lower performance, hence in the main analysis we are really improving on the quality of the return predictions and are not timing volatility.

2.6.2 Sensitivity analysis

The use of different performance criteria can lead to different results. To test the robustness of the results, we compute three alternative performance measures: the modified Sharpe ratio (Gregoriou and Gueyie, 2003; Eling and Schuhmacher, 2007), the Manipulation-proof Performance Measure (MPPM) (Goetzmann et al., 2007), and maximum drawdown (MDD). The modified Sharpe ratio divides the realized return by Value-at-Risk modified for skewness and kurtosis, as compared to the variance as in the Sharpe ratio $ModSR_s = \frac{\overline{r_s - r_f}}{MVAR_s}$, where $MVAR_s = -[r_s + \sigma_s(z_\alpha + (z_\alpha^2 - 1)\frac{Skew_s}{6} + (z_\alpha^3 - 3z_\alpha)\frac{Kurt_s}{24} - (2z_\alpha^3 - 5z_\alpha)\frac{Skew^2}{36})]$ and $z_\alpha = -1.96$. Goetzmann et al. (2007) show that performance measures of active management can be manipulated by managers. They propose a manipulation-proof measure, which accounts for non-linear payoffs in the return-risk relation. The measure is robust to the distribution of portfolio returns and does not require the assumption of a particular utility function to rank portfolios, like the performance fee. The third measure, maximum drawdown, is the maximum cumulative loss from the strategys peak wealth (cumulated return) to the following trough: $MDD_s = \min(X_1, X_2, \ldots, X_T)$, where $X_i = \sum_{j=1}^{i} r_{k+j}$, $1 \leq i \leq T$ and $1 \leq k \leq T$. Table 2.A.4 in the Appendix shows that the Zeros measure outperforms all the other strategies using the alternative evaluation criteria, e.g. the Zeros maximum drawdown is $8-16\%$ smaller than ILR.

All previous results are based on target conditional volatility $\sigma_s^* = 10\%$ and Relative Risk Aversion (RRA) $\delta = 5$. Tables 2.A.5 and 2.A.6 in the Appendix provide some sensitivity analysis. Table 2.A.5 in the Appendix shows the results for target conditional volatility equal to 15% and 20%. The SRs decrease slightly but the ranking of the strategies remains the same. The switching fees between the strategies are larger, in absolute value, because the strategies are more volatile, which a risk-averse investor dislikes.

The sensitivity results of the RRA parameter δ are also presented in Table 2.A.6 in the Appendix. The RRA is set to $\delta = 1$ in Panel A and $\delta = 10$ in Panel B. Similar RRA values are used in Fleming, Kirby, and Ostdiek (2001) and Della-Corte, Sarno, and Tsiakas (2009). The target conditional volatility is $\sigma_s^* = 10\%$ as in the main analysis.

When an investor is less risk averse, Panel A, she is willing to pay a higher switching fee to switch to the high return Zeros strategy. For example, the switching fee from the Roll strategy to the Zeros strategy in column (7) increases to 706.9 basis points per year from 480 in Table 2.3. When an investor is more risk averse, Panel B, she favors less volatile strategies. The ranking based on economic value remains the same: Zeros performs the best, followed by ILR, Roll, Effective Tick, and High-Low.

2.6.3 Bias adjustment

It is possible that market microstructure noise affects the analysis of observed returns, see Asparouhova, Bessembinder, and Kalcheva (2010, 2012) . They argue that "*Temporary deviations of trade prices from fundamental values impart bias to estimates of mean returns to individual securities, to differences in mean returns across (equally weighted) portfolios, and to parameters estimated in return regressions.*" To correct for this bias we can either weigh returns by their prior gross return (return-weighted) or by their prior firm value (value-weighted). The return-weighted approach places equal weights on all securities while the value-weighted approach gives more weight to large firms. Since our stocks are equally weighted within the portfolios, the natural choice is the return-weighted approach. Table 2.A.7 in the Appendix shows that the main results remain unchanged when stocks within the portfolios are return weighted instead of equally weighted.

2.7 Conclusion

In this paper we examine which proxy a liquidity timer should use. We build on the findings of Amihud (2002), Jones (2002), Baker and Stein (2004), and Bekaert, Harvey, and Lundblad (2007) who show that liquidity is predictable and that liquidity significantly predicts future excess returns. We investigate liquidity timing by measuring the economic value of liquidity forecasts from different liquidity proxies for investors, who engage in short-horizon asset allocation strategies. The following five low-frequency liquidity measures are considered in our liquidity timing analysis: illiquidity ratio (ILR) (Amihud, 2002), Roll (Roll, 1984), Effective Tick (Holden, 2009; Goyenko, Holden, and Trzcinka, 2009), Zeros (Lesmond, Ogden, and Trzcinka, 1999), and High-Low (Corwin and Schultz, 2012).

In line with Cao, Chen, Liang, and Lo (2013), we find that liquidity timing leads to tangible economic gains. The best performing strategy is based on the Zeros measure of Lesmond, Ogden, and Trzcinka (1999). Its Sharpe ratio is 0.51, over the sample period January 1947 - December 2008. The positive switching fees indicate that a risk-inverse investor will pay a high performance fee to switch from a strategy based on the ILR, Roll, Effective Tick, or High-Low measure to the Zeros strategy. The performance of the liquidity strategies is not driven by an alternative return predictor that is correlated with liquidity. Furthermore the performance of the strategies is not related to a certain subperiod and the ranking based on economic value is robust to different specifications and parameter settings.

The Zeros measure outperforms the other liquidity measures due to its robustness. It achieves positive performance even when the market is going down. The performance of the Zeros measure decreases when restricting the asset allocation weights and shrinking the return predictions. This implies that also the most extreme weights of the Zeros strategy are based on return predictions that are in the right direction.

Our ranking based on economic value differs from the ranking of Goyenko, Holden, and Trzcinka (2009), which is based on statistical criteria. They find that Effective Tick is good in measuring effective and realized spread and ILR is good in measuring price impact. We find that Zeros is the best proxy to use for liquidity timing in contrast to Effective Tick, which seems not relevant for predicting excess returns. This implies that low frequency measures that are a good proxy for high frequency transaction costs do not necessarily lead to the highest economic value in a liquidity timing strategy.

2.A Robustness results

Table 2.A.1 CRSP sample characteristics

The table shows the daily sample characteristics. Price is the stock price in \$. Volume is daily trading dollar volume in \$ millions, Market cap. is the market capitalization in \$ millions, Spread is the bid-ask spread, ask price−bid price in \$, Rel. Spread is Spread/((ask + bid)/2) in %, ILR is the illiquidity ratio |return|/dollar volume for a million shares, Volatility is the annualized daily standard deviation of returns, turnover is the annualized ration of volume/shares outstanding.

	Price	Volume	Market cap.	Spread	Rel. spread	ILR	Volatility	Turnover
Mean	31.0	9.65	2109	0.287	0.012	0.5471	24%	74%
Median	25.5	0.25	252	0.250	0.008	0.0189	16%	30%
25th	15.9	0.04	65	0.040	0.001	0.0007	7%	12%
75th	38.9	2.40	1010	0.375	0.017	0.1902	32%	78%
St. dev.	27.6	52.91	10332	0.518	0.013	12.1633	28%	168%

Table 2.A.2 Performance of dynamic asset allocation strategy based on LM1 liquidity measure

The table presents dynamic strategies performance for the LM1 liquidity measure (Liu, 2006). The dynamic asset allocation strategies are based on a mean-variance framework where the investor maximizes the conditional expected return subject to a target conditional volatility ($\sigma_s^* = 10\%$). 10Y are the return predictions of the risky assets based on a 10-year rolling window and Comb. are based on the average prediction of a 5, 10, and 20 year rolling window. The weights of the risky assets are not restricted. The sample period is January 1947 to December 2008, and the strategies start trading in January 1967. All numbers are annualized, except for the break-even costs τ_1 and τ_A that are reported in basis points per trade. The switching fee is the maximum performance fee a risk-averse investor is willing to pay to switch from one strategy to another. It is expressed in annual basis points and is computed based on a relative risk aversion parameter of 5. τ_1 shows the break-even costs under the assumption that all assets have the same transaction costs. τ_A shows the break-even costs in terms of large firms, whereby the liquidity differences between small and large firms are taken into account.

		Switching fee								
	SR	ILR	Eff. tick	Roll	Zeros	High-Low	τ_1	τ_A	Excess return	Volatility
10Y	0.41	337.7	547.5	530.2	50.6	687.2	4.3	2.0	4.64%	11.17%
Comb.	0.44	177.0	655.3	422.3	−87.9	632.6	4.8	2.2	5.25%	12.07%

Table 2.A.3 Benchmark results

The table presents the performance of the benchmark strategies. These strategies do not condition on liquidity information to obtain investment weights. The strategies in Panel A and B condition on past return information using either a 10 year rolling window or a combination of three rolling windows, respectively. The strategies in Panel C solve a different optimization problem to obtain the risky asset weights. The Equally weighted strategy gives all available stocks the same weight. Volatility timing minimizes the conditional expected portfolio variance subject to a target conditional return ($\mu_s^* = 10\%$), i.e. $\min_{w_t} \left\{ \sigma^2_{s,t+1|1} = w_t' \Sigma_{t+1|t} w_t \right\}$ s.t. $\mu_s^* = w_t' r_{k,t+1|t} + (1 - w_t' \mathbf{1}) r_{f,t}$. The "return predictions" $\left(r_{k,t+1|t}\right)$ are set equal to their unconditional averages, such that the weights only depend on the conditional covariance matrix. The minimum variance portfolio finds the portfolio with the lowest possible variance, again the investment weights only depend on the conditional covariance matrix. The weights of the risky assets are in all strategies unrestricted and all strategies start trading in January 1967.

	SR	τ_1	τ_A	Excess return	Volatility
Panel A: 10 year window					
Prevailing mean	0.15	6.2	2.8	1.74%	11.63%
Lagged return	0.18	2.0	0.9	2.19%	12.46%
Panel B: Combination of windows					
Prevailing mean	0.06	3.2	1.5	0.76%	11.97%
Lagged return	0.17	2.0	0.9	2.07%	12.49%
Panel C: Alternative optimization function					
Equally weighted	0.28	454.2	207.2	5.14%	18.30%
Volatility timing	0.16	15.1	6.8	1.88%	12.05%
Minimum variance	0.15	15.8	7.5	2.37%	15.36%

Table 2.A.4 Alternative performance measures

The table presents additional performance measures of the dynamic strategy results for the different liquidity measures. The dynamic asset allocation strategies are based on a mean-variance framework where the investor maximizes the conditional expected return subject to a target conditional volatility ($\sigma_s^* = 10\%$). The return predictions of the risky assets are based either on a 10-year rolling window, Panel A, or on the average prediction of a 5, 10, and 20 year rolling window, Panel B. The weights of the risky assets are not restricted. The sample period is January 1947 to December 2008, and the strategies start trading in January 1967. The liquidity measures are: ILR (Amihud, 2002), Roll (Roll, 1984), Effective Tick (Holden, 2009; Goyenko et al., 2009), Zeros (Lesmond et al., 1999), and HighLow (Corwin and Schultz, 2012). The modified Sharpe ratio (Mod. SR) of Gregoriou and Gueyie (2003) divides the expected return by Value-at-Risk modified for skewness and kurtosis. The Manipulation-proof Performance Measure (MPPM) shows the strategies premium return after adjusting for risk and can be interpreted as the annualized continuously compounded excess return certainty equivalent of the strategy (Goetzmann et al., 2007). The maximum drawdown (MDD) is the maximum cumulative loss from the strategys peak to the following trough.

	Mod. SR	MPPM	MDD
Panel A: 10 year window			
ILR	0.08	0.08%	−52.6%
Eff. tick	−0.03	−2.09%	−74.9%
Roll	−0.02	−1.86%	−66.2%
Zeros	0.27	3.00%	−36.6%
High-low	−0.08	−3.38%	−70.8%
Panel B: Combination of windows			
ILR	0.17	1.76%	−36.9%
Eff. tick	−0.06	−2.95%	−75.9%
Roll	0.04	−0.63%	−52.2%
Zeros	0.42	4.47%	−28.7%
High-low	−0.05	−2.63%	−57.6%

Table 2.A.5 Different target conditional volatility

The table presents dynamic strategy results for the different liquidity measures. The dynamic asset allocation strategies are based on a mean-variance framework where the investor maximizes the conditional expected return subject to a target conditional volatility (σ_s^*). In Panel A $\sigma_s^* = 15\%$ and in Panel B $\sigma_s^* = 20\%$. The return predictions of the risky assets are based either on a 10-year rolling window or on the average prediction of a 5, 10, and 20 year rolling window. The weights of the risky assets are not restricted. The sample period is January 1947 to December 2008, and the strategies start trading in January 1967. The liquidity measures are: ILR (Amihud, 2002), Roll (Roll, 1984), Effective Tick (Holden, 2009; Goyenko et al., 2009), Zeros (Lesmond et al., 1999), and High-Low (Corwin and Schultz, 2012). All numbers are annual, except for the break-even costs τ_1 and τ_A that are reported in basis points per trade. The switching fee is the maximum performance fee a risk-averse investor is willing to pay to switch from one strategy to another. It is expressed in annual basis points and is computed based on a relative risk aversion parameter of 5. τ_1 shows the break-even costs under the assumption that all assets have the same transaction costs. τ_A shows the break-even costs in terms of large firms, whereby the liquidity differences between small and large firms are taken into account. If the excess return of a strategy is negative we do not compute break-even transaction costs and report the symbol "-".

		Switching fee							
	SR	ILR	Eff. tick	Roll	Zeros	τ_1	τ_A	Excess return	Volatility
Panel A: Target conditional volatility $\sigma_s^* = 15\%$									
10 year window									
ILR	0.10					1.1	0.5	1.72%	17.81%
Eff. tick	−0.10	−288.3				–	–	−1.67%	17.12%
Roll	−0.07	−283.1	5.6			–	–	−1.25%	17.65%
Zeros	0.35	422.8	712.5	706.9		3.9	1.9	6.32%	18.08%
High-low	−0.19	−530.6	−240.9	−247.5	−953.4	–	–	−3.43%	17.99%
Combination of windows									
ILR	0.24					2.5	1.1	4.23%	17.31%
Eff. tick	−0.16	−729.4				–	–	−2.83%	17.71%
Roll	0.03	−374.5	353.7			0.5	0.2	0.57%	17.54%
Zeros	0.49	386.2	1114.7	759.4		5.2	2.4	8.62%	17.69%
High-low	−0.12	−711.6	18.5	−335.6	−1098.0	–	–	−2.28%	18.23%
Panel B: Target conditional volatility $\sigma_s^* = 20\%$									
10 year window									
ILR	0.07					0.8	0.4	1.57%	23.75%
Eff. tick	−0.13	−357.0				–	–	−2.88%	22.82%
Roll	−0.10	−371.3	−13.6			–	–	−2.35%	23.53%
Zeros	0.32	552.7	912.2	925.3		3.6	1.7	7.73%	24.11%
High-low	−0.22	−714.4	−355.8	−342.2	−1267.5	–	–	−5.25%	23.99%
Combination of windows									
ILR	0.22					2.2	1.0	4.97%	23.08%
Eff. tick	−0.19	−988.4				–	–	−4.44%	23.61%
Roll	0.00	−508.8	478.6			0.2	0.1	0.07%	23.38%
Zeros	0.46	497.8	1485.9	1007.3		4.8	2.3	10.89%	23.59%
High-low	−0.15	−987.9	4.7	−476.5	−1486.9	–	–	−3.76%	24.31%

Table 2.A.6 Different relative risk aversion parameter

The table presents the switching fees of the dynamic strategies based on the different liquidity measures. The dynamic asset allocation strategies are based on a mean-variance framework where the investor maximizes the conditional expected return subject to a target conditional volatility ($\sigma_s^* = 10\%$). The weights of the risky assets are not restricted. The sample period is January 1947 through December 2008, and the strategies start trading in January 1967. The liquidity measures are: ILR (Amihud, 2002), Roll (Roll, 1984), Effective Tick (Holden, 2009; Goyenko et al., 2009), Zeros (Lesmond et al., 1999), and High-Low (Corwin and Schultz, 2012). The switching fee is the maximum performance fee a risk-averse investor is willing to pay to switch from one strategy to another. It is expressed in annual basis points and requires a value for the relative risk aversion (RRA) parameter (in the main results it is set equal to 5). In Panel A the RRA is equal to 1 and in Panel B it is equal to 10.

	Switching fee			
	ILR	Eff. tick	Roll	Zeros
Panel A: RRA parameter $\delta = 1$				
10 year window				
ILR				
Eff. tick	−227.8			
Roll	−197.3	30.5		
Zeros	296.0	523.9	493.6	
High-low	−344.5	−116.6	−146.9	−640.8
Combination of windows				
ILR				
Eff. tick	−465.1			
Roll	−238.0	226.9		
Zeros	276.6	742.0	514.9	
High-low	−425.9	39.3	−187.8	−703.1
Panel B: RRA parameter $\delta = 10$				
10 year window				
ILR				
Eff. tick	−178.6			
Roll	−187.5	−7.5		
Zeros	275.6	455.6	463.1	
High-low	−358.1	−177.7	−171.6	−633.8
Combination of windows				
ILR				
Eff. tick	−494.1			
Roll	−254.1	239.1		
Zeros	249.4	743.0	503.9	
High-low	−494.1	2.3	−240.0	−743.4

Table 2.A.7 Strategy performance based on bias adjustment

This table shows the counterpart of the main results table. The only difference is that all individual stocks are weighted (within the 10 size portfolios) by their previous month gross return, instead of equally weighted. This adjustment is suggested in Asparouhova, Bessembinder, and Kalcheva (2013) to deal with noisy security prices.

		Switching fee							
	SR	ILR	Eff. Tick	Roll	Zeros	τ_1	τ_A	Excess return	Volatility
Panel A: 10 year window									
ILR	0.08					0.9	0.4	0.96%	11.88%
Eff. tick	0.02	−31.6				0.3	0.2	0.27%	11.08%
Roll	−0.02	−113.2	−80.6			–	–	−0.24%	11.73%
Zeros	0.33	286.4	318.7	400.3		3.7	1.8	3.98%	12.03%
High-low	−0.15	−267.2	−234.4	−150.0	−553.1	–	–	−1.74%	11.78%
Panel B: Combination of windows									
ILR	0.23					2.3	1.1	2.66%	11.61%
Eff. tick	−0.08	−348.3				–	–	−0.87%	11.57%
Roll	0.11	−139.9	208.4			1.2	0.6	1.28%	11.69%
Zeros	0.45	247.3	595.8	387.2		4.8	2.3	5.27%	11.69%
High-low	−0.08	−371.7	−23.4	−230.6	−619.0	–	–	−0.98%	11.83%

Chapter 3

Exploiting commodity momentum along the futures curves

Based on de Groot, Karstanje, and Zhou (2014)

3.1 Introduction

Several studies document a cross-sectional momentum effect in commodity futures markets. Erb and Harvey (2006) report a return of more than 10% per annum on a portfolio that longs commodity futures with the highest prior 12-month returns and shorts the worst-performing commodity futures. Miffre and Rallis (2007) extend this strategy for different ranking and holding periods up to 12 months and find profitable results for almost all definitions. Shen, Szakmary, and Sharma (2007) also report highly significant positive returns for holding periods up to nine months. In addition, Pirrong (2005) and Asness, Moskowitz, and Pedersen (2009) investigate momentum in multiple asset classes including commodities. What these commodity studies have in common is that only the nearest futures contracts are used for both the construction and implementation of momentum signals. Often futures contracts of various maturities are available for a given commodity. By considering only the nearest futures contract, the majority of investable deferred futures is not considered. This collection of futures could potentially offer additional information and investment opportunities.[1] We

[1] Various theories exist that try to explain the shape of the commodities futures curve. The oldest is the Normal Backwardation theory of Keynes (1930). Cootner (1960, 1967) generalizes the Normal Backwardation theory into the Generalized Hedging Pressure theory, while Kaldor (1939) and Working (1948, 1949) introduce an alternative explanation named the Theory of Storage.

propose alternative cross-sectional momentum strategies utilizing information further along the futures curve. We demonstrate that these strategies perform significantly better than a traditional momentum strategy.[2]

We identify four reasons why the futures curve potentially offers valuable information when exploiting a momentum strategy: contracts further along the curve could (i) exhibit more attractive roll yields, (ii) exhibit lower volatility, (iii) expand the opportunity set of our investable universe and (iv) lower the turnover of the portfolios. We will elaborate on these possible advantages in more detail. First, the excess returns of commodity futures can be decomposed in spot and roll returns, where roll return is defined as the yield that an investor captures when the futures price converges to the spot price as the futures contract comes closer to expiration, assuming that the spot price does not change.[3] The standard approach of investing in the nearest contracts might not be optimal in capturing roll returns. Commodity index providers have noticed the possible adverse effects of roll returns because long-only investments suffer from negative roll returns when the futures curve is upward sloping, i.e. is in contango. Miffre (2012) shows that long-only indices developed to minimize the exposure of negative roll returns have performed better than traditional long-only indices which are rolled based on the nearest contracts. Mouakhar and Roberge (2010) investigate the added value of maximizing the roll yield of long-only investments compared to simply buying the nearest contract in each of ten individual commodity futures. They find that buying the futures contract with the largest expected roll yield, as measured by the lowest price slope between two consecutive maturities, adds a return of on average 4.8% per year on top of buying the nearest futures contract. So far, this strand of literature has focused on enhancing traditional (long-only) indices and on stand-alone roll-yield strategies. However, it is not clear whether there is also added value to achieve on top of active momentum strategies.

Second, besides the possibility of finding more attractive roll yields, Samuelson (1965) argues that the volatility of futures returns decreases when the maturity of contracts increases. An economic argument is that most supply and demand shocks occur at the front-end of the

[2] A related stream of literature investigates so-called time-series as opposed to cross-sectional momentum strategies, see e.g. Szakmary et al. (2010), Moskowitz et al. (2012) and Baltas and Kosowski (2013). The main difference is that these time-series strategies construct commodity portfolios with possibly more long than short positions or vice versa, which implies that part of the strategy consists of commodity market timing. In our research, we focus on the cross-sectional pure momentum strategies without any market timing.

[3] This is under the assumption that the shape of the futures curve does not change. Note that it is difficult to ex-post decompose excess returns into spot and roll returns since both the level and the shape of the curve might have changed.

curve. Hence the prices of these front contracts react most heavily to news, while prices further along the curve are influenced less as there is more time to overcome the shocks. Daal et al. (2006) investigate this maturity effect empirically using an extensive futures dataset. They find that the effect tends to be stronger in agricultural and energy commodities than in financial futures. A possible implication of this maturity effect is that the volatility of a momentum strategy could be reduced by investing in futures with a longer maturity.

Third, even for the same commodity, contracts with different maturities exhibit large differences in returns and risks. For example in our data we find for lean hogs an average annualized return of -6.2% for the first contract, compared to 4.8% for the fifth contract. For WTI crude oil, we see an average annualized volatility of 33.2% for the front contract, compared to 22.2% for the tenth contract. These findings illustrate that non-front contracts behave differently from front contracts and essentially represent different investment opportunities. Therefore just like including more commodities into the universe, including non-front contracts further down the futures curves is expected to expand the opportunity set of our investable universe, which could potentially lead to more refined choices of contracts and better investment results.

And fourth, an interesting feature of buying contracts further along the curve is that these can potentially be kept longer in the portfolio. Contracts bought at the front-part of the curve soon need to be traded to avoid delivery, even though the commodity is still found to be attractive. On the other hand, as the trading volumes of contracts further on the curve are lower on average, the costs for trading a contract at the back-end of the curve could potentially be higher.

To exploit these four possible benefits, we propose three alternative momentum strategies in which we integrate term-structure information when generating and implementing momentum signals. All three strategies aim to reduce volatility by trading further on the curve and furthermore specifically aim to capture one or more of the above mentioned possible advantages. As a benchmark we take a cross-sectional generic momentum strategy that each month longs the commodities with the highest past 12-month returns (winner commodities) and shorts those with the lowest past 12-month returns (loser commodities).

The first alternative strategy that we propose aims to take advantage of the first benefit by maximizing the roll yield. More precisely, for the winner commodities we buy the most backwardated contract on the futures curve and for the loser commodities we sell the most contangoed contract, where we only include futures contracts that expire within 12 months. We show that implementing this roll-yield strategy on top of a traditional long-short momentum strategy generates significantly higher risk-adjusted returns, as the Sharpe ratio increases by more than 30% to 0.96 compared to 0.73 for the traditional front-contract momentum strategy. The improvement is both due to lower risk and higher returns.

The second strategy that we propose expands the traditional cross-sectional momentum strategy with curve momentum information. For each commodity, we first select the contract on the curve with the strongest and weakest momentum. We then cross-sectionally rank the commodities according to the selected contracts and long (short) the contracts with the highest (lowest) momentum. Besides enlarging our investment opportunity set, we implicitly take roll information into account as, even when a parallel shift in the term structure occurs, differences in roll return can cause differences in momentum returns along the curve.[4] We find that incorporating curve momentum leads to significantly higher returns (Sharpe ratios) compared to a traditional momentum strategy, namely 14.48% (0.97) versus 11.43% (0.73).

Our third strategy aims for higher roll returns and a much lower turnover compared to a traditional momentum strategy. We examine a strategy that remains invested in a particular contract even though it might not have the most optimal roll yield anymore. Only when the contract is about to expire or when the commodity switches from the long to the short portfolio (or vice versa) we again determine the most optimal contract. We observe that applying this strategy leads to a reduction in turnover of more than 50% compared to a traditional momentum strategy.

To ensure that the excess returns are not absorbed by transaction costs, we examine the added value that is created when the momentum strategies are actually implemented. Although transaction costs in futures markets are considerably lower compared to stocks, the turnover of momentum strategies is relatively high, which means that the impact of costs could still be substantial. Therefore, we incorporate two different trading cost schemes based on estimates of Szakmary, Shen, and Sharma (2010). Additionally, we contribute to the literature on commodity trading costs by proposing a third transaction cost scheme that links transaction costs to liquidity.[5] This ensures that transaction costs are higher for less liquid contracts, a component not covered by existing transaction cost schemes. We find that for all

[4]Momentum returns are based on excess futures returns, which are a combination of changes in the spot price and the roll yield.

[5]We thank an anonymous referee for this useful suggestion.

alternative momentum strategies and under all assumptions for transaction costs, alternative momentum strategies deliver higher returns and Sharpe ratios than for the generic momentum strategy. For example, using conservative trading cost estimates of approximately 22 basis points per trade, we observe that net returns increase from an insignificant 3.98% per annum for a traditional momentum strategy up to an economically and statistically significant 8.42% annual return for our alternative momentum strategies.

We next investigate whether the stronger returns of the alternative momentum strategies can be attributed to implicitly loading on the commodity market factor or on the carry strategy. The carry strategy takes long positions in the most backwardated (or least contangoed) commodities and short positions in the most contangoed (or least backwardated) commodities, see e.g. Erb and Harvey (2006) and Gorton and Rouwenhorst (2006). To examine this we regress the returns of the alternative momentum strategies on possible explanatory factor returns. We find economically and statistically significant alphas and therefore it is unlikely that our results are driven by implicit loadings on the market or carry factor. Fuertes et al. (2010) examine a double-sorted strategy of momentum and carry and find that buying the backwardated winners and shorting the contangoed losers outperforms a single momentum or carry strategy. Our results differ from this study as our alternative momentum strategies have added value beyond the momentum and carry factors. To further strengthen the finding that our approach adds value on top of well-known factors, we show that our proposed alternatives can also be profitably applied on top of such a double-sort strategy.

Finally, we analyze whether the additional profits of the momentum strategies that incorporate term-structure information are a compensation for lower liquidity. Besides imposing liquidity-dependent trading costs, we therefore perform a series of analyses to investigate this hypothesis in more detail. First, we examine whether the additional profits are due to investing in the back-end of the curve, where liquidity might be lowest. More specifically, we reduce the maximum maturity of futures contracts from 12 to 6 months and conclude that the additional profits are not driven by investing in futures contracts at the back-end of the curve. Second, we examine the impact of liquidity on our results more directly, by evaluating the momentum strategies when excluding the least liquid futures contracts from our universe. By using two types of liquidity measures, namely dollar trading volume and the Amihud (2002) illiquidity measure, we observe that the additional profits remain large and significant. Third, if we use a one-day implementation lag to ensure there is enough time to implement the trades, both gross and net performances remain similar. And fourth, we conclude that integrating term-structure information in momentum strategies also has substantial added value from 2000 onward, when more investors participated in commodity markets and overall liquidity was the largest. Hence, we conclude that it is unlikely that the additional profits are a compensation for lower liquidity.

The remainder of this article is organized as follows. We start in Section 3.2 by describing the data and analyzing the futures risk and return characteristics. We describe the four momentum strategies and the methodology to estimate transaction costs in Section 3.3. In Section 3.4 we present our main results and the portfolio return regressions. In Section 3.5 we show the results of several liquidity analyses. Section 3.6 presents our conclusion.

3.2 Data

Our investment universe consists of the constituents of the Standard & Poors Goldman Sachs Commodity Index (S&P GSCI) over the period January 1990 to September 2011.[6] We start with 18 commodity series at the beginning of our sample; all 24 series are available from July 1997. The sample includes six energy commodities (Brent crude oil, West Texas Intermediate crude oil, gasoil, heating oil, natural gas and RBOB gasoline); seven metals (gold, silver, aluminum, copper, lead, nickel and zinc); four softs (cocoa, coffee, cotton, and sugar); four grains (corn, soybeans, Chicago wheat and Kansas City wheat); and three meat commodities (feeder cattle, lean hogs and live cattle). We follow the S&P GSCI methodology and use data from the futures primary exchange, as the futures contracts of some of these commodities trade on multiple exchanges.[7] Furthermore, we only examine the individual futures contracts included in the S&P GSCI.[8] The number of distinct contracts a year varies per commodity; e.g. all the energy and industrial metal commodities have 12 distinct contracts a year, while cotton and sugar only have four distinct contracts. In addition to the

[6]Before 1990 the number of futures contracts diminishes quite rapidly.

[7]The Brent crude oil, gasoil, cocoa, coffee, cotton, and sugar data are from the Intercontinental Exchange (ICE); the West Texas Intermediate crude oil, heating oil, natural gas, and RBOB gasoline data are from the New York Mercantile Exchange (NYMEX); the gold and silver data are from the Commodity Exchange, Inc. (COMEX); the aluminum, copper, lead, nickel, and zinc data are from the London Metals Exchange (LME); the corn, soybeans, and Chicago wheat data are from the Chicago Board of Trade (CBOT); the Kansas wheat data are from the Kansas Board of Trade (KBT); and the feeder cattle, lean hogs, and live cattle data are from the Chicago Mercantile Exchange (CME).

[8]See Table 1 in the 2013 S&P GSCI Methodology for the selected 2013 futures contracts (http://www.spindices.com/documents/methodologies/methodology-sp-gsci.pdf).

above selection criteria we follow Mouakhar and Roberge (2010) and only include futures contracts in our analyses that expire within 12 months. For all individual contracts we collect futures prices from Bloomberg. Consistent with a large body of commodity research, such as Bessembinder (1992), Erb and Harvey (2006) and Miffre and Rallis (2007), we assume that the investment is made on a fully-collateralized basis.[9] In this case, the total monthly return of the investor is the change in month-end settlement prices plus the risk-free interest rate (e.g. the U.S. T-bill rate) earned from the deposit account. In our study, we focus on the changes in settlement prices, which we refer to as excess returns similar to Gorton, Hayashi, and Rouwenhorst (2013).

Table 3.1 reports the annualized excess returns, volatilities, average dollar trading volumes and the Amihud illiquidity measure of all commodity futures over our sample period. When we consider Panel A, we observe a large dispersion in average returns across commodities. For example, for the nearest contracts we find the lowest average return of -16.2% per annum for natural gas and the highest return of 11.8% per annum for gasoline. This indicates the potential benefits of correctly predicting which commodities to invest in. Moreover, we also find large return differences along the futures curve, although these are somewhat smaller on average. For example, for lean hogs we observe an annualized return of -6.2% for the first contract and 4.8% for the fifth contract. These return differences support our idea of enhancing a traditional momentum strategy by selecting the optimal contract on the curve. From these numbers we can also conclude that contracts along the same futures curve are not perfectly correlated with each other. Therefore, the inclusion of non-front contracts into the investable universe is likely to expand the opportunity set of the strategies and lead to better results.

[9] The advantage of assuming a fully-cash-collateralized investment is threefold. First, the investment process is largely simplified as there will be no leveraged positions which require extra deposit in or withdrawal from the margin account from time to time. Second, the calculation of the real-world return is fairly straight-forward, and no longer depends on the assumption of the initial margin. Third, the investment results are then presented in the most conservative manner, as strategy performances based on leverage are typically inflated compared to the base case which is fully-cash-collateralized.

Table 3.1 Summary statistics

This table presents the annualized excess returns (Panel A), volatilities (Panel B), average monthly dollar trading volumes (Panel C) and Amihud (2002) illiquidity measures (Panel D) of the 24 commodity futures from the nearest contract (i.e. first contract) up to the furthest contract with a maximum maturity of 12 months. The sample period is from January 1990 to September 2011. The trading volumes are computed as number of contracts traded multiplied by contract size multiplied by contract price and are expressed in million dollars. The Amihud (2002) illiquidity measure is computed as the monthly average of absolute daily return divided by the daily dollar trading volume.

	Xth nearest contract									
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Panel A: Return										
Energy										
Brent crude oil	10.0%	11.4%	11.2%	10.6%	11.5%	11.4%	11.4%	14.6%	14.4%	–
WTI crude oil	5.9%	8.6%	9.7%	10.0%	10.2%	10.4%	10.2%	10.0%	9.8%	9.5%
Gasoil	8.1%	8.0%	7.7%	8.6%	9.2%	8.3%	8.6%	10.5%	9.8%	9.3%
Heating oil	5.5%	6.7%	8.3%	9.1%	9.2%	9.3%	9.5%	9.5%	9.7%	9.8%
Natural gas	16.2%	8.0%	2.9%	1.3%	1.3%	0.3%	0.1%	1.4%	3.8%	4.9%
Gasoline	11.8%	11.7%	11.8%	12.1%	11.7%	11.7%	11.1%	9.3%	9.5%	9.6%
Metals										
Gold	2.9%	2.9%	2.9%	2.8%	–	–	–	–	–	–
Silver	4.2%	4.6%	4.8%	4.9%	–	–	–	–	–	–
Aluminum	2.5%	0.2%	0.4%	0.1%	0.6%	0.7%	0.9%	1.2%	1.5%	1.6%
Copper	8.1%	10.4%	10.0%	10.5%	10.8%	11.0%	11.2%	11.3%	11.4%	11.5%
Lead	8.2%	10.3%	9.7%	9.8%	9.8%	9.8%	9.6%	10.2%	10.4%	12.3%
Nickel	7.9%	10.4%	10.2%	10.9%	11.3%	11.4%	11.5%	11.7%	11.7%	11.7%
Zinc	3.5%	0.1%	0.0%	0.5%	1.0%	1.4%	1.9%	2.2%	2.6%	0.4%
Softs										
Cocoa	4.2%	2.6%	2.0%	1.6%	–	–	–	–	–	–
Coffee	3.8%	2.7%	2.7%	2.2%	–	–	–	–	–	–
Cotton	3.6%	1.1%	0.4%	–	–	–	–	–	–	–
Sugar	4.6%	5.6%	5.5%	–	–	–	–	–	–	–
Grains										
Corn	6.7%	3.8%	2.6%	1.4%	–	–	–	–	–	–
Soybeans	1.9%	3.2%	2.1%	2.8%	–	–	–	–	–	–
Chicago wheat	8.3%	4.3%	2.0%	1.5%	–	–	–	–	–	–
Kansas wheat	2.1%	0.7%	1.3%	1.6%	–	–	–	–	–	–
Meats										
Feeder cattle	2.0%	3.8%	4.4%	4.4%	4.2%	3.6%	–	–	–	–
Lean hogs	6.2%	4.5%	4.3%	4.3%	4.8%	–	–	–	–	–
Live cattle	0.4%	3.4%	2.1%	2.3%	2.4%	–	–	–	–	–
Panel B: Volatility										
Energy										
Brent crude oil	31.3%	29.4%	28.5%	27.3%	24.2%	23.5%	22.8%	23.1%	22.6%	–
WTI crude oil	33.2%	30.9%	29.2%	27.8%	26.5%	25.4%	24.4%	23.5%	22.8%	22.2%
Gasoil	31.9%	30.2%	28.8%	27.5%	26.6%	25.1%	24.4%	24.0%	22.5%	22.1%

Continued on next page

	Xth nearest contract									
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Heating oil	32.1%	30.3%	28.9%	27.6%	26.5%	25.4%	24.5%	23.8%	23.2%	22.7%
Natural gas	51.0%	43.6%	37.4%	33.7%	31.4%	29.2%	27.3%	25.6%	24.7%	24.0%
Gasoline	34.4%	31.1%	28.8%	27.3%	26.0%	25.0%	24.4%	24.4%	24.4%	23.4%
Metals										
Gold	15.5%	15.5%	15.5%	15.4%	–	–	–	–	–	–
Silver	28.3%	28.2%	28.0%	27.9%	–	–	–	–	–	–
Aluminum	20.0%	19.7%	19.4%	19.1%	18.8%	18.6%	18.3%	18.1%	17.9%	17.6%
Copper	28.2%	27.9%	27.8%	27.7%	27.4%	27.2%	27.0%	26.7%	26.5%	26.3%
Lead	31.5%	30.8%	30.5%	30.2%	29.8%	29.5%	29.4%	29.2%	29.1%	29.7%
Nickel	37.9%	37.7%	37.4%	37.0%	36.4%	35.8%	35.3%	34.8%	34.4%	34.0%
Zinc	28.5%	28.2%	27.9%	27.6%	27.3%	27.1%	26.9%	26.7%	26.5%	24.7%
Softs										
Cocoa	30.1%	29.1%	28.1%	27.3%	–	–	–	–	–	–
Coffee	38.5%	35.8%	34.0%	32.7%	–	–	–	–	–	–
Cotton	26.9%	25.0%	22.7%	–	–	–	–	–	–	–
Sugar	31.6%	28.4%	25.6%	–	–	–	–	–	–	–
Grains										
Corn	25.4%	24.5%	23.1%	21.6%	–	–	–	–	–	–
Soybeans	23.5%	22.8%	21.9%	20.9%	–	–	–	–	–	–
Chicago wheat	27.6%	26.3%	24.5%	22.3%	–	–	–	–	–	–
Kansas wheat	26.9%	25.8%	24.6%	22.7%	–	–	–	–	–	–
Meats										
Feeder cattle	13.0%	12.1%	11.0%	10.5%	10.0%	9.7%	–	–	–	–
Lean hogs	23.7%	20.2%	16.9%	14.7%	13.7%	–	–	–	–	–
Live cattle	12.9%	10.5%	8.9%	8.1%	8.0%	–	–	–	–	–
Panel C: Trading Volume (expressed in million dollars)										
Energy										
Brent crude oil	1505.7	641.7	324.7	211.9	151.4	113.2	97.9	86.1	65.2	–
WTI crude oil	5091.5	1734.0	780.3	435.2	295.9	218.0	164.9	139.8	112.0	89.6
Gasoil	854.4	344.2	172.7	107.4	79.0	59.5	43.4	41.3	34.0	24.7
Heating oil	943.5	331.4	171.1	104.8	73.2	55.8	40.6	31.8	20.8	14.5
Natural gas	1038.3	436.5	250.4	163.7	124.6	99.0	79.9	68.8	54.0	43.5
Gasoline	952.4	352.6	170.2	91.3	53.5	32.7	19.4	13.1	9.2	6.3
Metals										
Gold	2830.1	113.9	38.4	22.4	–	–	–	–	–	–
Silver	868.7	49.3	17.2	7.2	–	–	–	–	–	–
Aluminum	553.2	521.9	174.6	86.6	71.5	57.0	46.6	40.8	29.5	30.6
Copper	676.8	626.2	190.1	112.5	86.4	67.3	49.9	54.3	43.2	37.7
Lead	77.0	52.5	18.9	11.4	8.8	6.6	7.5	7.5	5.7	4.8
Nickel	143.9	133.8	40.0	16.6	13.0	8.2	7.1	5.6	5.0	4.4
Zinc	195.9	176.8	52.5	23.2	24.6	18.4	11.2	11.4	6.8	6.7
Softs										
Cocoa	96.8	24.1	22.1	4.1	–	–	–	–	–	–
Coffee	317.7	70.9	22.1	9.4	–	–	–	–	–	–

Continued on next page

	Xth nearest contract									
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Cotton	246.7	75.7	28.9	–	–	–	–	–	–	–
Sugar	325.2	115.6	48.9	–	–	–	–	–	–	–
Grains										
Corn	932.8	380.9	170.3	85.8	–	–	–	–	–	–
Soybeans	1667.5	361.6	158.6	80.3	–	–	–	–	–	–
Chicago wheat	509.7	158.3	61.5	29.7	–	–	–	–	–	–
Kansas wheat	149.4	49.5	19.5	9.3	–	–	–	–	–	–
Meats										
Feeder cattle	82.3	17.9	8.6	3.7	1.3	0.6	–	–	–	–
Lean hogs	161.9	59.2	26.3	12.0	5.6	–	–	–	–	–
Live cattle	315.5	130.9	55.3	23.7	7.3	–	–	–	–	–
Panel D: Amihud illiquidity measure (expressed in basis points per one-million-dollar trade)										
Energy										
Brent crude oil	0.9	5.5	18.7	37.1	71.0	149.3	120.2	155.3	229.3	–
WTI crude oil	0.6	1.5	1.5	3.6	3.2	12.3	13.0	23.4	38.5	87.8
Gasoil	1.3	4.1	18.3	45.2	95.6	137.0	192.4	159.3	305.2	243.0
Heating oil	4.4	4.2	6.6	8.1	13.8	37.3	82.3	122.1	172.9	295.5
Natural gas	5.8	7.2	8.8	16.3	21.5	25.8	36.2	48.4	60.6	85.5
Gasoline	2.6	3.8	6.0	19.6	70.3	159.9	341.6	409.9	555.6	777.8
Metals										
Gold	6.5	43.0	105.5	159.4	–	–	–	–	–	–
Silver	6.9	105.7	378.0	629.4	–	–	–	–	–	–
Aluminum	3.1	5.0	8.5	9.8	9.7	17.1	18.6	21.5	40.6	35.1
Copper	1.8	2.9	5.0	10.9	9.2	11.2	21.9	20.4	24.0	31.0
Lead	25.8	32.4	70.4	124.4	124.7	214.3	238.5	231.4	246.2	245.5
Nickel	13.8	10.5	20.1	63.2	74.1	121.9	172.3	199.5	253.3	237.5
Zinc	13.8	19.4	18.7	39.3	48.3	86.9	154.8	143.4	224.5	159.0
Softs										
Cocoa	56.7	55.1	99.2	433.7	–	–	–	–	–	–
Coffee	22.7	46.9	72.5	140.6	–	–	–	–	–	–
Cotton	21.2	17.8	36.5	–	–	–	–	–	–	–
Sugar	26.3	19.4	25.6	–	–	–	–	–	–	–
Grains										
Corn	0.3	1.1	2.6	9.0	–	–	–	–	–	–
Soybeans	0.5	8.3	14.9	68.5	–	–	–	–	–	–
Chicago wheat	0.8	6.7	24.6	157.8	–	–	–	–	–	–
Kansas wheat	1.5	6.5	65.8	364.5	–	–	–	–	–	–
Meats										
Feeder cattle	1.8	6.3	15.1	39.4	113.1	238.0	–	–	–	–
Lean hogs	1.1	3.2	8.9	28.7	95.1	–	–	–	–	–
Live cattle	0.3	0.6	1.3	4.4	19.4	–	–	–	–	–

Besides return differences we observe large differences in volatilities from Panel B of Table 3.1. For WTI crude oil, we see an annualized volatility of 33.2% for the front contract, compared to 22.2% for the tenth contract. In line with Samuelson (1965) we find that in almost all cases, volatility decreases when the time to maturity increases. Hence, strategies that trade in more distant contracts could potentially exhibit a lower volatility.

In Panel C we present average trading volume in million dollars, computed by multiplying the number of contracts traded by the contract size, and then multiplying this by the price in dollars.[10] We observe large differences in this liquidity measure among commodities. For example, the trading volume of crude oil is much higher than that of lead. In addition we also observe large differences along the curve, as e.g. the first contract of Brent oil has an average trading volume of $1,506$ million dollar, while that of the ninth contract is 65 million dollar. That more distant futures are less often traded than nearby contracts confirms that most investors use nearby contracts to take positions.

To give more insight in liquidity differences across different contracts, we also compute the Amihud (2002) illiquidity measure, which measures the price impact of a trade. It is computed as the monthly average of daily absolute return divided by dollar trading volume. We multiply the measure with one million, so that Panel D shows the return impact in basis points of a one-million-dollar trade. The results in Panel D are in line with the trading volume results. In general, the Amihud measure increases further along the curve, i.e. a one-million-dollar trade has a larger impact on more distant futures prices and hence those contracts are less liquid. Across commodities there are also large differences, e.g. a one-million-dollar trade in the first WTI crude oil contract impacts prices by 0.6 basis points on average, while a trade of the same size in the first cocoa contract influences prices by 56.7 basis points.

Overall, the variation of average returns and volatilities along each commodity futures curve indicate the potential added value of integrating term-structure information into a generic momentum strategy. However, liquidity measures indicate that more distant contracts are less often traded, so there will be a trade-off between the improvement in performance and the increase in trading costs.

[10]The data for the number of contracts traded are from Bloomberg. For industrial metals these data are available from around 2005. We therefore approximate dollar volume by deflating the volume in January 2005 back in time by 9.8% per annum, which is the average annual change in dollar volume of the available commodity futures contracts from 1990 to 2004.

3.3 Methodology

3.3.1 Constructing momentum strategies

To investigate commodity momentum strategies that integrate term-structure information, we construct four different types of momentum portfolios. The first portfolio is a "generic momentum" strategy that represents the traditional momentum strategy documented by Erb and Harvey (2006). Each month-end, we rank all commodities cross-sectionally according to the past 12-month returns of their nearest contracts. This portfolio takes long (short) positions in the 50% of commodities with the highest (lowest) returns, using equal weights.[11] We then compute the return of this portfolio in the following month.

Our first alternative strategy is an optimal-roll momentum portfolio, where we aim to maximize the roll yield. Compared to the generic momentum strategy, we select the same commodities for the long and short portfolios. However, this portfolio does not necessarily invest only in the front contracts as is the case for the generic strategy. Instead, for the 50% of commodities with the most attractive returns, we long the contract on the curve with the largest price slope (the most backwardated or least contangoed). The slope of contract i is defined as

$$\frac{f_t^{i-1} - f_t^i}{f_t^i \left(\tau_i - \tau_{i-1}\right)}, \tag{3.1}$$

where f_t^i is the futures price of contract i at time t with time to maturity τ_i and f_t^{i-1} is the futures price of the adjacent contract with time to maturity τ_{i-1}.[12] For the 50% of commodities with the least attractive returns, we short the contract on the curve with the smallest slope (the most contangoed or least backwardated).

The second alternative strategy is an "all-contracts momentum" portfolio, where we expand a traditional cross-sectional momentum strategy with curve momentum information.

[11] There seems to be little consistency in the literature on the construction of commodities portfolios. Both Erb and Harvey (2006) and Gorton and Rouwenhorst (2006) construct top and bottom 50% portfolios. Miffre and Rallis (2007) investigate the top/bottom 20%, while Shen et al. (2007) the top/bottom 33%. Unreported results show that the returns of more concentrated portfolios go up. However, we simultaneously observe that the volatility of these portfolios increase even more, which results in portfolios with lower Sharpe ratios. The number of assets in commodity portfolios becomes very small once we move to the top and bottom 20%, as the commodity-specific risk of the portfolios increases.

[12] To compute the slope corresponding to an investable first contract we extrapolate the futures curve using a piecewise cubic interpolation method, see Fritsch and Carlson (1980). The advantage of this method is that it preserves the shape of the data and respects monotonicity. This method ensures we can also invest in the most nearby contract.

For this portfolio, each month we compute the 12-month return of all contracts along the curve for all commodities.[13] We then select for each commodity the contract with the highest 12-month return and flag it as a candidate for the long portfolio. Likewise, we also select for each commodity the contract with the lowest 12-month return and label it as a short candidate. After repeating this selection process for all commodities, we next rank all contracts indicated as candidates for the long portfolio and long the 50% of commodities with the highest momentum. Similarly, we rank all short candidates and short the 50% of commodities with the lowest momentum. What is different from the generic and optimal-roll strategies is that this portfolio might take both long and short positions in different contracts of the same commodity, a situation which might occur when there is a large dispersion in the momentum values of the contracts for a particular commodity.

As portfolio turnover, and therefore trading costs, is relatively high for typical momentum strategies, we in addition examine a third alternative momentum strategy, using a "low-turnover roll momentum" portfolio. With this portfolio we still aim for a higher return due to better roll positions, but with a much lower turnover compared to the other momentum strategies. Contracts bought on the front-end of the curve regularly need to be traded, as these contracts are the closest to expiration. Even if according to the strategy a commodity remains in the portfolio, the position in its nearest contract still needs to be replaced (i.e. rolled forward) after a short period of time. An advantage of buying contracts further along the curve is that these could be kept in the portfolio for much longer. Compared to the optimal-roll momentum strategy where each month we determine the contracts with the most optimal slope, we now remain invested in the same contract unless it is about to expire, or the commodity changes from the long to the short portfolio (or vice versa) based on its front-contract momentum. In that case, we take a new position in the contract with the most optimal slope. This way, we will not always have positions in the most optimal contracts and therefore expect a lower gross return of this strategy compared to the standard optimal-roll momentum strategy. However, due to a lower turnover, the expected trading costs are also

[13] In line with generic momentum, the 12-month return of the xth contract is based on the past returns of the xth nearby contract

lower. The impact on net return is therefore a trade-off between the expected reduction in gross return and the lower trading costs.[14]

3.3.2 Incorporating transaction costs

Although transaction costs involved with commodity futures are relatively low (see e.g. Locke and Venkatesh, 1997) and taking short positions is not more complex than taking long positions, momentum strategies typically exhibit high turnover. We therefore also evaluate the returns of the momentum strategies when incorporating realistic trading costs. We use three different transaction cost schemes: two are based on Szakmary et al. (2010) which are labeled as standard and conservative, while we propose a third novel scheme that incorporates trading cost variation along the futures curve based on liquidity differences.

The standard transaction costs scheme consists of a fixed brokerage commission of USD 10 per contract and a bid-ask spread of one tick. Szakmary et al. (2010) estimate transaction costs (TC) as a percentage of the notional contract value in month t:

$$TC_t = [10 + (Ticksize \times CM)]/(Price_t \times CM), \tag{3.2}$$

where the tick size is measured in dollars, CM is the contract multiplier (i.e. the number of units of the underlying commodity deliverable per contract) and $Price_t$ is the price of the contract in dollars at the end of month t.[15] For conservative transaction costs, which might reflect the actual costs of large-scale trading activity, the brokerage commission is assumed to be USD 20 per contract and the bid-ask spread three ticks instead of one:

$$TC_t = [20 + 3 \times (Ticksize \times CM)]/(Price_t \times CM), \tag{3.3}$$

Compared to the standard cost estimates, the conservative estimates assume a market impact that is three times higher for trades in the same commodity futures contract, which is in line with the findings of Marshall et al. (2012). They conclude that a more aggressive trader

[14] Due to the construction of the all-contracts momentum strategy there is no low-turnover parallel for this strategy, as the contracts with the most extreme momentum are selected before, instead of after, the cross-sectional comparison is made.

[15] The tick size defines the minimum price movement of a futures contract. It varies across different commodities and is specified in the contract specifications of each futures contract. We retrieved all tick size and contract multiplier data from the futures exchange websites.

who requires immediate liquidity exhibits costs on average three times higher compared to a more patient trader who splits the futures trades over one hour. The standard and conservative trading cost estimates could therefore also be interpreted as the costs associated with patient and more aggressive trading styles respectively.

Front contracts are in general more liquid, which could potentially lead to lower trading costs than contracts on the back end of the curve. Unfortunately there is little information available on the relationship between trading costs and the time to maturity, as the academic literature in this area is scarce. To bridge this gap, we propose a third methodology, where we assume a linear relation between the Amihud illiquidity measure and transaction costs. Each month we assume for each commodity that the most liquid contract (i.e. with the lowest Amihud estimate) trades against the standard transaction costs while the least liquid contract with the highest Amihud estimate trades against the conservative costs. This implies that we assume that trading illiquid contracts is around three times more expensive. For the intermediate contracts we assume that the costs increase proportionally to the increase in the Amihud illiquidity measure. The implication of this trading costs scheme is that trading in contracts further on the curve involves higher trading costs.

We condition trading costs on the Amihud illiquidity measure because of two reasons. First, Marshall et al. (2012) find that the Amihud illiquidity measure has the largest correlation with high-frequency liquidity benchmarks. Second, besides a negative relation between trading volume and transaction costs, there is in general also a positive relation between volatility and transaction costs.[16] This implies that on the one hand estimated trading costs for more distant futures are higher due to their lower trading volumes, while on the other hand lower due to their lower volatility. Therefore, the Amihud illiquidity measure is highly appropriate as it incorporates both volatility, in the form of absolute returns, and trading volume.

[16]In equities the positive relation between volatility and transaction costs is well established. For example. Chordia et al. (2001a) note it is well known that individual stock volatility is cross-sectionally associated with higher [bid-ask] spreads (Benston and Hagerman (1974). In commodities this relation between volatility and transaction costs is documented by Marshall et al. (2012). They only investigate front contracts and establish this relation based on time variation of volatility and transaction costs over time, and not along the futures curve.

3.4 Results

3.4.1 Profitability of strategies including term-structure information

In our first empirical analysis we evaluate the momentum profits of the generic momentum strategy and the strategies incorporating term-structure information. Panel A of Table 3.2 reports the average annualized gross returns and associated t-statistics, the volatilities and the Sharpe ratios of the generic momentum strategy and the three alternative momentum strategies. Furthermore, for all alternative strategies we show the Ledoit and Wolf (2008) test statistic, which evaluates whether the Sharpe ratios of the alternative strategies are significantly different from that of the generic momentum strategy. In this test we take into account the possibility that strategy returns can be non-normal and auto-correlated. As the distribution of this statistic is non-standard, the reported P-values are based on bootstrap resamples. In addition, the table also contains the maximum drawdown, the average maturity of the contracts deployed by each strategy and the average single-counted and one-sided annual turnover. This means that an annual turnover of 100% indicates that the long and short portfolio is completely changed once a year.

Table 3.2 Strategies performance − 12 months maturity bound

This table shows the performance of four cross-sectional commodity momentum strategies over the sample period January 1990 to September 2011. Panel A reports the gross annualized performance while Panel B, C and D report the net annualized performance based on respectively standard, Amihud-based and conservative trading costs. "Generic momentum" ranks commodities according to the past 12-month returns of the nearest contracts and longs (shorts) the 50% of commodities with the highest (lowest) returns. "Optimal-roll momentum" also ranks commodities based on front-contract momentum, but longs (shorts) the contract with the most backwardated (most contangoed) slope. "All-contracts momentum" first selects the contract on each commodity curve with the highest (lowest) 12-month returns and then longs (shorts) the 50% of commodities with the highest (lowest) momentum. "Low-turnover roll momentum" compared to optimal-roll momentum, which monthly determines the contracts with the most optimal slope, remains invested in the same contract unless it is about to expire or the commodity changes position. All portfolios are equally weighted. The Ledoit and Wolf (2008) test (LW-statistic) evaluates whether the Sharpe ratios of the alternative strategies are significantly different from that of the generic momentum strategy. The average maturity of the contracts in portfolio is presented in months. The turnover figures presented in this table are single-counted and one-sided. In addition the average single-trip costs of the transactions in basis points (bps) are shown.

	Generic	Optimal roll	All contracts	Low-turnover roll
Panel A: Gross returns				
Return	11.43%	13.05%	14.48%	12.31%
t-statistic	3.33	4.36	4.44	4.03
Volatility	15.65%	13.62%	14.86%	13.91%
Sharpe ratio	0.73	0.96	0.97	0.88
LW-statistic	–	2.67	3.29	2.67
p-value	–	0.01	0.00	0.01
Max. drawdown	−23.70%	−21.21%	−21.57%	−21.09%
Maturity (months)	1.50	5.01	3.85	4.22
Turnover	855%	756%	880%	402%
Panel B: Net returns assuming standard transaction costs				
Return	8.43%	10.19%	11.27%	10.76%
t-statistic	2.45	3.40	3.44	3.51
Volatility	15.69%	13.65%	14.91%	13.95%
Sharpe ratio	0.54	0.75	0.76	0.77
LW-statistic	–	2.55	3.06	4.14
p-value	–	0.01	0.00	0.00
Max. drawdown	−25.30%	−22.74%	−22.00%	−21.96%
Avg. costs (bps)	8.10	8.56	8.20	8.71
Panel C: Net returns assuming Amihud-based transaction costs				
Return	8.36%	9.01%	10.34%	10.32%
t-statistic	2.43	3.00	3.15	3.37
Volatility	15.70%	13.65%	14.94%	13.96%
Sharpe ratio	0.53	0.66	0.69	0.74
LW-statistic	–	1.56	2.25	3.63
p-value	–	0.12	0.02	0.00
Max. drawdown	−25.30%	−23.33%	−22.21%	−22.29%
Avg. costs (bps)	8.30	12.26	10.67	11.22
Panel D: Net returns assuming conservative transaction costs				
Return	3.98%	5.94%	6.52%	8.42%
t-statistic	1.15	1.98	1.98	2.74
Volatility	15.79%	13.71%	15.00%	14.01%
Sharpe ratio	0.25	0.43	0.43	0.60
LW-statistic	–	2.33	2.67	6.24
p-value	–	0.02	0.01	0.00
Max. drawdown	−30.59%	−25.75%	−23.55%	−23.24%
Avg. costs (bps)	20.95	22.17	21.24	22.29

Consistent with Erb and Harvey (2006) and Miffre and Rallis (2007), we find large and significant profits for the generic momentum strategy. More specifically, the strategy earns a gross annual return of 11.43% and a gross Sharpe ratio of 0.73. Although the results are strong, we observe even higher risk-adjusted returns for our alternative momentum strategies. The Sharpe ratios of the alternative strategies range between 0.88 for low-turnover roll momentum and 0.97 for all-contracts momentum, due to higher returns and lower volatilities. These lower portfolio volatilities are in line with Samuelson (1965), as the average maturities of the contracts range between 3.85 months for the all-contracts strategy and 5.01 months for the optimal-roll strategy, all above the average maturity of 1.50 months for the generic strategy. Note that the average maturities are well below 12 months, which is the maximum maturity of the contracts we invest in. This indicates that the strategies on average invest more in contracts on the front part of the curve, where liquidity is likely to be highest. The Sharpe ratios of the alternative momentum strategies are significantly different from the generic strategy as the P-values are lower than 0.05. Moreover, the maximum drawdowns of the alternative strategies are all smaller than that of the generic strategy.

In addition, we observe that the low-turnover roll momentum strategy lives up to its name as it exhibits a turnover that is approximately 50% lower than the other strategies. The turnover of 855% per annum for the generic strategy implies that a portfolio manager needs to completely change the portfolio every 1.4 months on average. The implications on net returns can be observed in Panel B, C, and D of Table 3.2, assuming respectively the Szakmary et al. (2010) standard trading cost estimates, our Amihud-dependent estimates and the conservative cost estimates. The table also presents the average single-trip costs of the transactions in basis points. We observe that trading costs have a significant impact on the return of momentum strategies. As the average trading costs using the standard cost estimates are around 8 basis points per trade, we observe a deterioration in return of around 3% for the high-turnover momentum strategies. Note that this average cost is substantially higher than the often used 3.3 basis points as reported in Locke and Venkatesh (1997) and used by among others Fuertes et al. (2010). The net Sharpe ratios of the optimal-roll and all-contracts momentum strategies of 0.75 and 0.76 respectively remain significantly higher than the net Sharpe ratio of 0.54 for the generic strategy. The impact of trading costs on the low-turnover roll strategy is much lower, resulting in only a 1.55% lower return. The net return and Sharpe ratio of this strategy are now higher than those of the optimal-roll momentum strategy.

Panel C reports the net results when assuming Amihud-dependent transaction costs. In line with our expectations we observe higher average costs for our alternative momentum strategies, which on average trade further on the curve. The costs for the generic momentum strategy are around 8 basis points, while those of the alternative strategies range between 10.67 and 12.26 basis points. Nevertheless, also using this trading costs scheme, we observe higher net returns and Sharpe ratios and less negative maximum drawdowns for all our alternative momentum strategies.

The contrast in net returns among the different strategies becomes even larger once the conservative cost estimates are taken into account, which might better reflect the actual costs in the case of large-scale trading activity or a more aggressive trading style. In Panel D we find that the average cost estimates are now about 22 basis points per trade. The return of the generic strategy drops by 65%, leading to an insignificant 3.98% annualized return. Obviously, the impact on return is also large for the two alternative high-turnover momentum strategies. However, we still observe economically and statistically significant returns of 5.94% and 6.52% respectively for the optimal-roll and all-contracts momentum strategies, with Sharpe ratios of 0.43. Interestingly, when assuming these relatively high trading costs we observe the highest net returns (8.42%) and Sharpe ratio (0.60) for the low-turnover roll momentum strategy.

We conclude that incorporating term-structure information in momentum strategies leads to significantly higher Sharpe ratios. When facing relatively high trading costs, it might be important to smartly reduce portfolio turnover to preserve the majority of returns.

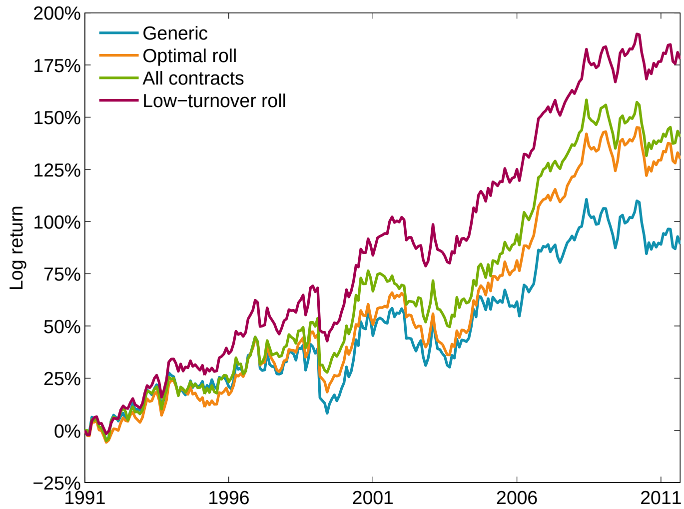
Finally, we analyze the returns of the different momentum strategies over time. Figure 3.1A shows the cumulative returns when incorporating conservative transaction costs of the four momentum strategies over our sample period. We observe that the generic momentum strategy obtains the lowest returns, while the highest returns are generated by our low-turnover roll momentum strategy. We also observe a gradual increase of the difference in returns, so that the added value of the strategies is not generated in one particular sub-sample period. Figure 3.1B presents 5-year rolling Sharpe ratios of all momentum strategies. We observe that during most 5-year sub-periods, the alternative momentum strategies obtain higher

Sharpe ratios than the generic momentum strategy. This confirms our previous finding that our results are not obtained during one particular sub-period in our sample.

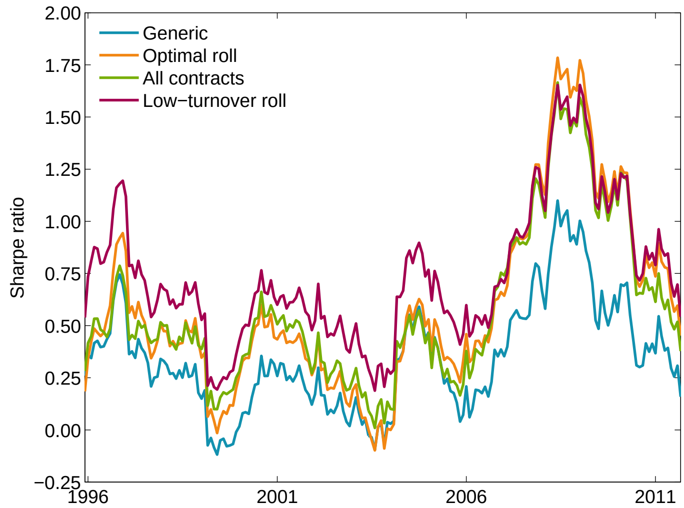
Figure 3.1 Strategies performance over time

These figures show the performance over time of our cross-sectional commodity momentum strategies over the sample period January 1991 - September 2011. The performance is based on net returns assuming conservative transaction costs. Subfigure A shows cumulative log returns and Subfigure B shows 5-year rolling Sharpe ratios. Note that subfigure B starts in 1996 due to the 5-year rolling window.

(a) Cumulative performance



(b) Rolling Sharpe Ratios



3.4.2 Portfolio return regressions

We continue our empirical analysis by investigating to what extent the profits of the momentum strategies that integrate term-structure information can be attributed to exposures to well-known commodity factor premiums. In particular we focus on the commodity market factor and the carry strategy. The carry strategy is based on term-structure information and takes long positions in the most backwardated commodities and short positions in the most contangoed commodities, see e.g. Erb and Harvey (2006) and Gorton and Rouwenhorst (2006). We regress the gross and net returns of the momentum strategies on a market and carry factor:

$$R_{i,t} = \alpha + \beta_1 Market_t + \beta_2 Carry_t + \varepsilon_{i,t} \tag{3.4}$$

where $R_{i,t}$ is the return of momentum strategy i in month t. $Market_t$ is the excess return of the commodity market index as proxied by the S&P GSCI in month t. $Carry_t$ is the return of a carry strategy in month t defined as an equally-weighted portfolio that longs (shorts) the 50% of commodities with the highest (lowest) annualized ratio of the nearest futures price to the next-nearest futures price. The coefficients α, β_1 and β_2 are to be estimated, and $\varepsilon_{i,t}$ is the residual return of strategy i in month t. In addition, we perform regression analyses where we add the generic momentum factor to analyze the added value of the alternative momentum strategies on top of the traditional momentum strategy. All coefficient estimates, associated t-statistics and R-squared values are presented in Table 3.3.

Table 3.3 Portfolio return regressions

This table presents the coefficient estimates, t-statistics (between brackets) and R-squared values obtained from regressions of the monthly gross (Panel A) and net returns using standard (Panel B), Amihud-based (Panel C) and conservative (Panel D) transaction costs of the four cross-sectional momentum strategies on the carry, market and momentum factors. Carry is the return of a strategy defined as an equally-weighted portfolio that longs (shorts) the 50% of commodities with the highest (lowest) annualized ratio of nearby futures price to the nearest futures price. Market is the excess return of the S&P GSCI market index. Momentum is the return of the generic momentum strategy. The intercepts of the regressions are annualized and reported as alpha.

	Generic	Optimal roll	All contracts	Low-turnover roll	Optimal roll	All contracts	Low-turnover roll
Panel A: Gross returns							
Alpha	5.17%	6.97%	8.29%	6.40%	2.84%	3.61%	2.04%
	(1.80)	(2.85)	(3.00)	(2.52)	(3.29)	(3.82)	(2.64)
Market	0.11	0.09	0.10	0.10	−0.00	0.00	0.00
	(2.93)	(2.71)	(2.80)	(2.88)	(−0.08)	(0.13)	(0.30)
Carry	0.66	0.60	0.61	0.60	0.07	0.01	0.04
	(10.82)	(11.57)	(10.43)	(11.03)	(3.35)	(0.61)	(1.94)
Momentum	–	–	–	–	0.80	0.90	0.84
	–	–	–	–	(42.09)	(43.29)	(49.27)
R^2	34.0%	36.6%	32.4%	34.8%	92.3%	92.2%	94.0%
Panel B: Net returns assuming standard transaction costs							
Alpha	4.25%	6.08%	7.11%	6.66%	2.68%	3.26%	3.08%
	(1.50)	(2.51)	(2.60)	(2.65)	(3.14)	(3.49)	(4.03)
Market	0.11	0.09	0.10	0.10	−0.00	0.00	0.00
	(2.94)	(2.72)	(2.81)	(2.88)	(−0.09)	(0.14)	(0.27)
Carry	0.66	0.60	0.62	0.60	0.07	0.02	0.04
	(10.90)	(11.59)	(10.52)	(11.05)	(3.22)	(0.65)	(1.80)
Momentum	–	–	–	–	0.80	0.90	0.84
	–	–	–	–	(42.00)	(43.28)	(49.43)
R^2	34.3%	36.7%	32.7%	34.9%	92.3%	92.2%	94.1%
Panel C: Net returns assuming Amihud-based transaction costs							
Alpha	4.22%	5.02%	6.29%	6.29%	1.64%	2.46%	2.73%
	(1.49)	(2.07)	(2.30)	(2.50)	(1.92)	(2.63)	(3.57)
Market	0.11	0.09	0.10	0.10	−0.00	0.00	0.00
	(2.94)	(2.73)	(2.82)	(2.89)	(−0.08)	(0.15)	(0.28)
Carry	0.66	0.60	0.62	0.60	0.07	0.02	0.04
	(10.91)	(11.58)	(10.53)	(11.06)	(3.18)	(0.66)	(1.80)
Momentum	–	–	–	–	0.80	0.91	0.84
	–	–	–	–	(41.89)	(43.24)	(49.50)
R^2	34.4%	36.7%	32.8%	34.9%	92.2%	92.2%	94.1%
Panel D: Net returns assuming conservative transaction costs							
Alpha	2.85%	4.67%	5.31%	7.04%	2.39%	2.74%	4.63%
	(1.01)	(1.94)	(1.96)	(2.82)	(2.82)	(2.96)	(6.12)
Market	0.11	0.09	0.10	0.10	−0.00	0.00	0.00
	(2.96)	(2.73)	(2.83)	(2.88)	(−0.10)	(0.14)	(0.22)
Carry	0.67	0.60	0.62	0.59	0.07	0.02	0.03
	(11.04)	(11.63)	(10.68)	(11.10)	(3.00)	(0.70)	(1.56)
Momentum	–	–	–	–	0.80	0.90	0.84
	–	–	–	–	(41.69)	(43.24)	(49.36)
R^2	34.9%	36.9%	33.4%	35.1%	92.2%	92.3%	94.1%

Panel A reports the results for the gross returns. For the generic momentum strategy we observe a large and significant exposure to the carry factor and also a significant exposure to the market factor. This leads to an annualized alpha of 5.17% with a t-statistic of 1.80 for the generic strategy compared to the 11.43% "raw" return in Table 3.2. The high coefficient of the carry factor is consistent with Gorton et al. (2013) who argue that momentum portfolios take positions in similar commodities to the carry-sorted portfolios. We notice that the (untabulated) correlation between the returns of the carry and generic momentum strategy is 0.56 over our sample period, in line with these findings. When we consider the regressions in the left part of the panel, we observe similar exposures to the market and carry factors for the alternative momentum strategies. The alphas of all momentum strategies with integrated term-structure information remain significant and are all larger than the alpha of generic momentum, ranging from 6.40% for the low-turnover roll momentum strategy to 8.29% for the all-contracts momentum strategy.

We next consider the regressions including the generic momentum factor in the right part of Panel A. For the alternative momentum strategies, we observe small and insignificant coefficient estimates for the market factor. Not surprisingly, we find large positive coefficient estimates for the momentum factor, which also explains the high explanatory power of the regressions with R-squared values above 90%. Due to the positive correlation between the carry and generic momentum factors, we now find much smaller coefficient estimates for the carry factor. Interestingly, the alphas remain significantly different from zero, ranging from 2.04% for the low-turnover roll momentum strategy to 3.61% for the all-contracts momentum strategy.

We next consider the results when regressing net returns using standard, Amihud-dependent and conservative trading cost estimates in respectively Panel B, C and D of Table 3.3. We evaluate these returns also against the net returns of the carry and momentum factors. The alphas of the optimal-roll and all-contracts momentum strategies become lower, but remain in almost all cases significant and larger than the insignificant net alphas of the generic strategy. On the contrary, the low-turnover roll momentum strategy earns for two out of three applied trading cost schemes a higher alpha. This results in an alpha of 7.04% when incorporating conservative trading costs and regressed against the market and carry factors, which is much larger than the 2.85% alpha of the generic momentum strategy.

We conclude that the returns of the alternative momentum strategies are not driven by exposures to well-known commodity factors.[17] Finally, untabulated results further indicate

[17] In addition, we also investigate to what extent the profits of the cross-sectional momentum strategies can be attributed to exposures to the time-series momentum factor of Moskowitz et al. (2012) and the Fung and Hsieh (2004) seven factor model. We conclude that the additional returns of the alternative momentum strategies cannot be explained by loading on the time-series momentum factor or on Fung and Hsiehs factors. All results are available on request.

that the optimal-roll and all-contracts momentum strategies are different types of strategies as their alphas have only a modestly positive correlation of 0.35.

3.4.3 Double-sort on momentum and carry

In this section we use an alternative method to examine the added value of our alternative strategies on top of momentum and carry strategies. Fuertes et al. (2010) show that double-sort strategies on both momentum and carry lead to superior results compared to a generic momentum strategy. We therefore apply our alternative strategies of integrating term-structure information on top of these double-sorts.

Our starting point is a generic double-sort strategy where we first sort 50% of the commodities into a winner and 50% in a loser portfolio, based on their past 12-month momentum. Next, we sort 50% of the commodities within this winner portfolio in a high- and 50% in a low-carry portfolio, where carry is defined as the front-contract slope in Formula (3.1). Also for the loser portfolio we apply this sort on carry. We then take equally-weighted long positions in the winner/high-carry commodities and short positions in the loser/low-carry commodities. This set-up is comparable to the double-sort investigated by Fuertes et al. (2010).[18]

The alternative strategies we propose are constructed in a similar fashion as described in Section 3.3.1. The optimal-roll double-sort is based on the same commodities as the generic double-sort, however does not necessarily invest only in front contracts. The long positions in the winner/high-carry portfolio are taken in the most backwardated contracts along the curve, while the short positions in the loser/low-carry portfolio are invested in the most contangoed contracts along the curve. The all-contracts strategy sorts on optimal 12-month momentum candidates instead of front contract momentum values, where optimal momentum candidates are determined in the same way as before for our single-sort momentum portfolios. The 50% highest (lowest) momentum candidates end up in the winner (loser) portfolio and within this winner (loser) portfolio we select the 50% commodities with the highest (lowest) carry. The strategy then longs winner/high-carry commodities and shorts loser/low-carry commodities. Our low-turnover roll alternative is a reduced turnover version

[18]The number of commodities in the long and short portfolio (six from 1998 onwards) is similar to Fuertes et al. (2010), who construct two momentum and three carry portfolios, but have more commodities in their universe.

of the first alternative strategy, similar to the single-sort version. The results of the double-sort strategies are presented in Table 3.4. As the number of commodities in the double-sort portfolios is smaller, namely six in the top and six in the bottom instead of two times twelve, we also construct a generic single-sort momentum strategy based on quartiles portfolios as a comparison. We long the 25% most attractive commodities and short the 25% least attractive commodities. The (untabulated) annualized returns of this strategy are 13.02% with a volatility of 24.88%. Compared to the generic portfolio in Table 3.2 we observe that a more concentrated portfolio generates a higher return, but also exhibits higher risk, which results in a lower Sharpe ratio (0.52 versus 0.73).

Table 3.4 Double-sort strategies performance

This table shows the performance of double-sort cross-sectional strategies on momentum and carry over the sample period January 1990 to September 2011. We sort first on momentum and then on carry. The alternative strategies are constructed in a similar fashion as in Table 3.2. Panel A reports the gross annualized performance while Panel B, C and D report the net annualized performance based on respectively standard, Amihud-based and conservative trading costs.

	Generic	Optimal roll	All contracts	Low-turnover roll
Panel A: Gross returns				
Return	16.09%	17.27%	19.85%	16.52%
t-statistic	3.43	3.99	4.33	3.78
Volatility	21.35%	19.71%	20.87%	19.90%
Sharpe ratio	0.75	0.88	0.95	0.83
LW-statistic	–	1.61	5.08	1.30
p-value	–	0.11	0.00	0.20
Max. drawdown	−26.19%	−27.28%	−26.15%	−27.62%
Maturity (months)	1.49	3.88	3.01	3.62
Turnover	926%	857%	933%	620%
Panel B: Net returns assuming standard transaction costs				
Return	12.60%	13.82%	16.19%	13.93%
t-statistic	2.68	3.19	3.52	3.18
Volatility	21.40%	19.76%	20.93%	19.95%
Sharpe ratio	0.59	0.70	0.77	0.70
LW-statistic	–	1.49	4.94	1.96
p-value	–	0.14	0.00	0.05
Max. drawdown	−32.56%	−28.75%	−27.67%	−28.81%
Avg. costs (bps)	8.38	8.85	8.46	9.15
Panel C: Net returns assuming Amihud-based transaction costs				
Return	12.50%	12.95%	15.59%	13.43%
t-statistic	2.66	2.98	3.39	3.07
Volatility	21.41%	19.78%	20.94%	19.96%
Sharpe ratio	0.58	0.65	0.74	0.67
LW-statistic	–	0.94	4.37	1.57
p-value	–	0.35	0.00	0.12
Max. drawdown	−32.57%	−29.06%	−27.91%	−29.07%
Avg. costs (bps)	8.61	11.15	9.88	10.95
Panel D: Net returns assuming conservative transaction costs				
Return	7.45%	8.72%	10.79%	10.08%
t-statistic	1.58	2.00	2.34	2.29
Volatility	21.50%	19.84%	21.04%	20.04%
Sharpe ratio	0.35	0.44	0.51	0.50
LW-statistic	–	1.29	4.69	2.97
p-value	–	0.20	0.00	0.00
Max. drawdown	−42.66%	−37.37%	−37.67%	−34.03%
Avg. costs (bps)	21.71	22.94	21.93	23.57

We confirm the findings of Fuertes et al. (2010) as the generic double-sort achieves a higher return (16.09%) and a lower volatility (21.35%) than the generic momentum strategy with similar number of commodities. When we compare our three alternative strategies with the generic double-sort in Panel A of Table 3.4 we observe that all Sharpe ratios are higher, both due to higher returns and lower volatilities. And also when we observe the net results in Panel B to D, we observe that in all cases returns the Sharpe ratios for the alternative double-sort strategies are higher than the generic double-sort strategies. These results again confirm that our alternative strategies add value beyond momentum and carry factors.

3.5 Liquidity analyses

This section presents the results of a series of analyses to investigate whether the additional profits of the alternative momentum strategies are a compensation for lower liquidity. In Subsection 3.5.1 we examine the sensitivity of our results by limiting our universe to futures contracts with a maximum maturity of six months. In Subsection 3.5.2 we analyze the impact of excluding the most illiquid contracts. We investigate the implication of a one-day trading lag in Subsection 3.5.3. Finally, in Subsection 3.5.4 we present the results after 2000.

3.5.1 Implementation with futures contracts up to six months maturity

We continue our empirical analyses by evaluating the alternative momentum strategies if we reduce the maximum maturity of futures contracts to invest in from 12 months to 6 months. All the other settings are exactly the same as with the main approach. Note that Table 3.2 indicates that the average maturities of the alternative momentum strategies are well below 12 months, as they range between 3.85 and 5.01 months. However, there could still be regular investments in the back-end of the curve. We perform this analysis to ensure that the additional profits of the momentum strategies that incorporate term-structure information are not due to investing in the back-end of the curve, where liquidity might be the lowest. The results are presented in Table 3.5.

Table 3.5 Strategies performance − six months maturity bound

This table shows the risk and return characteristics of four cross-sectional commodity momentum strategies over the sample period January 1990 to September 2011. The strategies are constructed in the same way as in Table 3.2 with the difference that the strategies invest in contracts with a maturity up to six months. Panel A reports the gross annualized performance while Panel B, C and D report the net annualized performance based on respectively standard, Amihud-based and conservative trading costs.

	Generic	Optimal roll	All contracts	Low-turnover roll
Panel A: Gross returns				
Return	11.43%	13.38%	13.47%	12.96%
t-statistic	3.33	4.23	4.01	3.96
Volatility	15.65%	14.40%	15.28%	14.89%
Sharpe ratio	0.73	0.93	0.88	0.87
LW-statistic	–	3.72	3.10	3.84
p-value	–	0.00	0.00	0.00
Max. drawdown	−23.70%	−21.66%	−21.09%	−22.19%
Maturity (months)	1.50	2.94	2.18	2.40
Turnover	855%	812%	853%	528%
Panel B: Net returns assuming standard transaction costs				
Return	8.43%	10.36%	10.43%	10.91%
t-statistic	2.45	3.27	3.10	3.33
Volatility	15.69%	14.41%	15.33%	14.93%
Sharpe ratio	0.54	0.72	0.68	0.73
LW-statistic	–	3.47	3.03	5.35
p-value	–	0.00	0.00	0.00
Max. drawdown	−25.30%	−22.85%	−21.47%	−22.98%
Avg. costs (bps)	8.10	8.44	8.07	8.76
Panel C: Net returns assuming Amihud-based transaction costs				
Return	7.89%	8.28%	9.30%	9.93%
t-statistic	2.29	2.62	2.76	3.02
Volatility	15.72%	14.42%	15.36%	14.96%
Sharpe ratio	0.50	0.57	0.61	0.66
LW-statistic	–	1.38	2.19	4.65
p-value	–	0.17	0.03	0.00
Max. drawdown	−25.45%	−24.04%	−22.26%	−23.34%
Avg. costs (bps)	9.60	14.50	11.18	13.02
Panel D: Net returns assuming conservative transaction costs				
Return	3.98%	5.87%	5.93%	7.82%
t-statistic	1.15	1.85	1.75	2.37
Volatility	15.79%	14.45%	15.42%	15.00%
Sharpe ratio	0.25	0.41	0.38	0.52
LW-statistic	–	3.00	2.90	7.37
p-value	–	0.00	0.00	0.00
Max. drawdown	−30.59%	−25.68%	−24.83%	−24.41%
Avg. costs (bps)	20.95	21.85	20.86	22.56

When we consider the gross returns in Panel A we observe that all three alternative momentum strategies remain able to deliver significantly higher risk-adjusted returns compared to a generic momentum strategy. When compared to the results in Table 3.2 with a 12-month maturity bound, we find slightly higher returns for the optimal-roll and low-turnover roll momentum strategies and somewhat lower returns for the all-contracts momentum strategy. As the average maturity of the contracts reduces by 1.5 to 2 months, the portfolios volatilities increase somewhat. This leads to slightly lower Sharpe ratios compared to the 12-month maturity-bound results.

We observe significant net returns for most of the alternative strategies. And even though the turnover of the low-turnover roll strategy increases from 402% to 528%, the strategy remains statistically significant when including conservative trading costs. Thus, we conclude that the additional profits are not driven by investing in futures contracts at the back-end of the curve.

3.5.2 Implementation on most liquid futures contracts

To more directly examine the impact of liquidity on our results, we next evaluate the momentum strategies when excluding the least liquid futures contracts from our universe. For this purpose we use two different types of liquidity measures, namely dollar trading volume and the Amihud illiquidity measure.

Each month in our sample period we exclude the most illiquid futures contracts according to a certain measure. This way we acknowledge that liquidity varies substantially across commodities, and that for less liquid commodities, more contracts will be excluded than for more liquid commodities. Assuming a USD 100 million long/short portfolio, we set the dollar volume trading threshold in such a way that we currently do not trade more than 25% of the trading volume of a particular contract. As the universe currently consists of 24 commodities, we have 12 long and short positions. The value of one trade is therefore USD $8.33\,(= 100/12)$ million, implying a dollar volume threshold of USD $33.33\,(= 8.33/0.25)$ million at the end of our sample period. We deflate this threshold back in time by 4.05%, which is the average annual total return of the S&P GSCI index during our sample period. As a result, we exclude almost 50% of the futures contracts from our universe. For the Amihud illiquidity measure, we set the threshold at 4 basis points at the end of our sample period, so that we also exclude about 50% of the most illiquid futures contracts from the universe. This threshold implies we exclude commodity futures contracts for which the price impact resulting from trading USD 1 million is more than 4 basis points. Similarly, the Amihud illiquidity threshold is inflated back in time by 4.05% per annum. The results of the four momentum strategies applied to the most liquid futures contracts based on dollar trading volume and the Amihud illiquidity measure are presented in Table 3.6 and 3.7 respectively.[19]

[19]Unreported analyses show that conclusions remain similar when different liquidity threshold values are used for both dollar trading volume and the Amihud illiquidity measure.

Table 3.6 Strategies performance − dollar trading volume screening

This table shows the risk and return characteristics of four cross-sectional commodity momentum strategies over the sample period January 1990 to September 2011. The strategies are constructed in the same way as in Table 3.2 with the difference that here we first exclude futures contracts which do not meet the minimum requirement of the monthly average of daily dollar trading volume. This threshold is set at USD 33.33 million at the end of our sample period and is deflated back in time by 4.05% per annum. Panel A reports the gross annualized performance while Panel B, C and D report the net annualized performance based on respectively standard, Amihud-based and conservative trading costs.

	Generic	Optimal roll	All contracts	Low-turnover roll
Panel A: Gross returns				
Return	11.90%	13.59%	15.55%	13.54%
t-statistic	3.42	4.26	4.54	4.14
Volatility	15.87%	14.54%	15.60%	14.89%
Sharpe ratio	0.75	0.93	1.00	0.91
LW-statistic	–	2.49	3.64	3.51
p-value	–	0.01	0.00	0.00
Max. drawdown	−24.84%	−22.46%	−22.13%	−22.40%
Maturity (months)	1.51	3.42	2.97	2.83
Turnover	860%	777%	838%	555%
Panel B: Net returns assuming standard transaction costs				
Return	8.90%	10.67%	12.51%	11.34%
t-statistic	2.55	3.33	3.65	3.46
Volatility	15.91%	14.58%	15.63%	14.92%
Sharpe ratio	0.56	0.73	0.80	0.76
LW-statistic	–	2.47	3.68	4.58
p-value	–	0.01	0.00	0.00
Max. drawdown	−26.79%	−24.43%	−22.52%	−23.92%
Avg. costs (bps)	8.02	8.48	8.06	8.89
Panel C: Net returns assuming Amihud-based transaction costs				
Return	8.81%	10.36%	12.24%	11.22%
t-statistic	2.52	3.23	3.57	3.42
Volatility	15.91%	14.59%	15.63%	14.93%
Sharpe ratio	0.55	0.71	0.78	0.75
LW-statistic	–	2.29	3.53	4.78
p-value	–	0.02	0.00	0.00
Max. drawdown	−26.81%	−24.65%	−22.61%	−24.04%
Avg. costs (bps)	8.26	9.41	8.80	9.42
Panel D: Net returns assuming conservative transaction costs				
Return	4.44%	6.33%	8.00%	8.04%
t-statistic	1.26	1.97	2.32	2.44
Volatility	15.99%	14.65%	15.69%	15.00%
Sharpe ratio	0.28	0.43	0.51	0.54
LW-statistic	–	2.40	3.72	6.22
p-value	–	0.02	0.00	0.00
Max. drawdown	−30.25%	−27.36%	−25.38%	−26.21%
Avg. costs (bps)	20.77	21.97	20.85	22.95

Table 3.7 Strategies performance − Amihud illiquidity measure screening

This table shows the risk and return characteristics of four cross-sectional commodity momentum strategies over the sample period January 1990 to September 2011. The strategies are constructed in the same way as in Table 3.2 with the difference that here we first exclude futures contracts which do not meet the maximum threshold of the Amihud illiquidity measure. This measure is calculated as the monthly average of the absolute daily return divided by the daily dollar trading volume. The maximum requirement is set at 4 basis points for a one-million-dollar trade at the end of our sample period and is inflated back in time by 4.05% per annum. Panel A reports the gross annualized performance while Panel B, C and D report the net annualized performance based on respectively standard, Amihud-based and conservative trading costs.

	Generic	Optimal roll	All contracts	Low-turnover roll
Panel A: Gross returns				
Return	12.02%	14.47%	15.39%	14.04%
t-statistic	3.51	4.51	4.52	4.32
Volatility	15.59%	14.61%	15.51%	14.81%
Sharpe ratio	0.77	0.99	0.99	0.95
LW-statistic	–	2.45	3.17	3.83
p-value	–	0.01	0.00	0.00
Max. drawdown	-23.72%	-22.10%	-22.06%	-22.05%
Maturity (months)	1.51	3.41	2.94	2.87
Turnover	868%	782%	844%	560%
Panel B: Net returns assuming standard transaction costs				
Return	8.98%	11.51%	12.29%	11.84%
t-statistic	2.62	3.58	3.61	3.64
Volatility	15.62%	14.63%	15.53%	14.83%
Sharpe ratio	0.57	0.79	0.79	0.80
LW-statistic	–	2.38	3.16	4.89
p-value	–	0.02	0.00	0.00
Max. drawdown	-24.88%	-23.31%	-22.46%	-23.25%
Avg. costs (bps)	8.05	8.48	8.17	8.79
Panel C: Net returns assuming Amihud-based transaction costs				
Return	8.85%	11.19%	12.02%	11.70%
t-statistic	2.58	3.48	3.53	3.59
Volatility	15.62%	14.63%	15.53%	14.83%
Sharpe ratio	0.57	0.76	0.77	0.79
LW-statistic	–	2.25	3.05	4.80
p-value	–	0.03	0.00	0.00
Max. drawdown	-24.92%	-23.50%	-22.72%	-23.37%
Avg. costs (bps)	8.40	9.43	8.90	9.34
Panel D: Net returns assuming conservative transaction costs				
Return	4.46%	7.11%	7.70%	8.53%
t-statistic	1.29	2.21	2.25	2.61
Volatility	15.69%	14.67%	15.57%	14.87%
Sharpe ratio	0.28	0.49	0.49	0.57
LW-statistic	–	2.41	3.13	6.18
p-value	–	0.02	0.00	0.00
Max. drawdown	-27.41%	-25.66%	-26.04%	-25.18%
Avg. costs (bps)	20.86	21.93	21.15	22.65

In Panel A of Table 3.6, we observe that the screening on dollar trading volume has a marginal impact on the gross performance of the four strategies. In all cases, we even observe a slightly higher gross return compared to the results in Table 3.2. Volatility increases as well, in line with the shorter average maturity of the invested contracts, which leads to similar gross Sharpe ratios. Results are similar when we apply the screening based on the Amihud illiquidity measure, as presented in Panel A of Table 3.7. As the screenings have hardly any impact on turnover, we also observe that the alternative strategies remain profitable after taking transaction costs into account (Panel B, C and D of Table 3.6 and 3.7). Only the turnover of the low-turnover roll momentum strategy increases from 402% per annum to 555% and 560% when we apply a screening on dollar trading volume and the Amihud illiquidity measure respectively. However, as the gross returns are also 1% to 2% higher for this strategy after applying a liquidity screening, the net returns and Sharpe ratios have a similar magnitude as the results in Table 3.2. In addition, we observe that all alternative momentum strategies significantly outperform the generic momentum strategy. We therefore conclude that the added value of incorporating term-structure information in momentum strategies is not due to investing in contracts with a low liquidity.

3.5.3 Results with a one-day implementation lag

We next examine the profitability of the momentum strategies assuming a one-day implementation lag. Although we use various trading costs estimates when evaluating the net returns of the strategies, it is still possible that their profitability is largely concentrated in the period just after rebalancing, and that the gross returns would decline significantly when there is a delay in trading. Marshall et al. (2012) claim that the commodity futures markets are resilient and that liquidity returns to pre-trade levels after 30-60 minutes. Therefore, we evaluate the profitability of the momentum strategies by assuming that investors have one full trading day to rebalance their portfolio and that by gradually implementing the new positions, the trade impact can largely be mitigated. More specifically, we construct the portfolios in a similar fashion as before, the difference being that we determine the long and short positions based on data up to the day before every month-end. The results are reported in Table 3.8.

Table 3.8 Strategies performance − one-day implementation lag

This table shows the risk and return characteristics of four cross-sectional commodity momentum strategies over the sample period January 1990 to September 2011. The strategies are constructed in a similar fashion as in Table 3.2 with the difference that the long and short positions are based on data up to the day before month-end. Panel A reports the gross annualized performance while Panel B, C and D report the net annualized performance based on respectively standard, Amihud-based and conservative trading costs.

	Generic	Optimal roll	All contracts	Low-turnover roll
Panel A: Gross returns				
Return	11.37%	12.79%	14.57%	12.23%
t-statistic	3.21	4.13	4.38	3.90
Volatility	16.12%	14.10%	15.16%	14.29%
Sharpe ratio	0.71	0.91	0.96	0.86
LW-statistic	–	2.84	3.44	2.57
p-value	–	0.00	0.00	0.01
Max. drawdown	−23.88%	−21.36%	−22.80%	−20.85%
Maturity (months)	1.49	5.01	3.85	4.21
Turnover	855%	768%	883%	400%
Panel B: Net returns assuming standard transaction costs				
Return	8.38%	9.91%	11.34%	10.72%
t-statistic	2.36	3.19	3.40	3.41
Volatility	16.16%	14.14%	15.21%	14.32%
Sharpe ratio	0.52	0.70	0.75	0.75
LW-statistic	–	2.60	3.14	4.03
p-value	–	0.01	0.00	0.00
Max. drawdown	−25.27%	−23.01%	−25.00%	−21.74%
Avg. costs (bps)	8.08	8.53	8.21	8.57
Panel C: Net returns assuming Amihud-based transaction costs				
Return	8.31%	8.68%	10.42%	10.27%
t-statistic	2.34	2.79	3.11	3.27
Volatility	16.17%	14.16%	15.24%	14.33%
Sharpe ratio	0.51	0.61	0.68	0.72
LW-statistic	–	1.36	2.37	3.54
p-value	–	0.18	0.02	0.00
Max. drawdown	−25.27%	−23.64%	−26.14%	−21.97%
Avg. costs (bps)	8.28	12.30	10.63	11.10
Panel D: Net returns assuming conservative transaction costs				
Return	3.94%	5.63%	6.57%	8.43%
t-statistic	1.11	1.80	1.96	2.67
Volatility	16.25%	14.22%	15.30%	14.38%
Sharpe ratio	0.24	0.40	0.43	0.59
LW-statistic	–	2.21	2.66	6.10
p-value	–	0.03	0.01	0.00
Max. drawdown	−29.05%	−25.46%	−28.89%	−23.05%
Avg. costs (bps)	20.91	22.08	21.25	21.93

When we consider the gross returns in Panel A, we observe only a slight reduction for most of the strategies compared to the returns without an implementation lag in Table 2. We therefore conclude that a one-day implementation has hardly any impact on our results.

3.5.4 Results since 2000

Due to the increased popularity of commodity investing since 2000, we conclude our empirical analyses by examining the momentum strategies from January 2000 to September 2011 when overall liquidity was the highest. The results are presented in Table 3.9. When we consider Panel A we observe higher gross returns for all strategies in the most recent 11 years of our sample compared to the returns of the whole sample period in Table 3.2. Next to that, we observe higher volatilities in the most recent sample period, in line with increased market volatility. All in all, we observe a similar Sharpe ratio for the generic momentum strategy and higher Sharpe ratios for the alternative momentum strategies. For example, the Sharpe ratio of the low-turnover roll momentum strategy is 0.99 during the last 11 years of our sample, while the ratio is 0.88 over the whole sample period.

We observe in Panel B, C and D of Table 3.9 that due to lower average costs since 2000, the differences in net returns are even larger. For instance, when assuming conservative transaction costs, the optimal-roll momentum strategy is able to deliver a higher net return of around 3%. Also in the recent period, the alternative strategies obtain higher returns at lower risk compared to the generic momentum strategy. We therefore conclude that integrating term-structure information in momentum strategies also has added value since 2000, when more investors participated in commodity markets and overall liquidity was the highest.

Table 3.9 Strategies performance after 2000

This table shows the risk and return characteristics of four cross-sectional commodity momentum strategies over the sample period January 2000 to September 2011. The strategies are constructed in the same way as in Table 3.9. Panel A reports the gross annualized performance while Panel B, C and D report the net annualized performance based on respectively standard, Amihud-based and conservative trading costs.

	Generic	Optimal roll	All contracts	Low-turnover roll
Panel A: Gross returns				
Return	11.97%	14.71%	15.23%	14.04%
t-statistic	2.54	3.62	3.42	3.39
Volatility	16.12%	13.92%	15.27%	14.21%
Sharpe ratio	0.74	1.06	1.00	0.99
LW-statistic	–	2.03	2.17	4.36
p-value	–	0.05	0.03	0.00
Max. drawdown	−21.39%	−19.62%	−21.57%	−19.03%
Maturity (months)	1.46	5.03	3.93	4.14
Turnover	881%	773%	893%	404%
Panel B: Net returns assuming standard transaction costs				
Return	9.58%	12.37%	12.68%	12.77%
t-statistic	2.03	3.04	2.84	3.07
Volatility	16.15%	13.94%	15.31%	14.24%
Sharpe ratio	0.59	0.89	0.83	0.90
LW-statistic	–	1.98	2.03	5.48
p-value	–	0.05	0.04	0.00
Max. drawdown	−21.78%	−20.01%	−22.00%	−19.19%
Avg. costs (bps)	6.17	6.75	6.33	6.98
Panel C: Net returns assuming Amihud-based transaction costs				
Return	9.54%	11.38%	11.92%	12.43%
t-statistic	2.03	2.80	2.66	2.99
Volatility	16.14%	13.95%	15.34%	14.25%
Sharpe ratio	0.59	0.82	0.78	0.87
LW-statistic	–	1.54	1.63	5.06
p-value	–	0.13	0.11	0.00
Max. drawdown	−21.78%	−21.28%	−22.21%	−19.35%
Avg. costs (bps)	6.28	9.68	8.28	8.83
Panel D: Net returns assuming conservative transaction costs				
Return	6.06%	8.89%	8.92%	10.85%
t-statistic	1.28	2.18	1.99	2.60
Volatility	16.19%	13.98%	15.38%	14.29%
Sharpe ratio	0.37	0.64	0.58	0.76
LW-statistic	–	1.89	1.83	7.10
p-value	–	0.06	0.07	0.00
Max. drawdown	−24.63%	−25.75%	−23.55%	−20.98%
Avg. costs (bps)	15.80	17.30	16.22	17.76

3.6 Conclusion

This study examines novel momentum strategies in commodities futures markets that incorporate term-structure information. Previous studies only use the nearest futures contracts both for the construction and implementation of momentum signals. These strategies might therefore potentially miss out on valuable information regarding the futures curve, such as the possibility that contracts further along the curve could exhibit more attractive roll yields and lower volatility.

We show that alternative momentum strategies which integrate term-structure information by selecting contracts on the curve with the largest expected roll-yield or with the strongest momentum earn significantly higher risk-adjusted returns than a traditional momentum strategy, even when incorporating three different transaction costs schemes.

To lower transaction costs even further, we examine another alternative momentum strategy aiming for higher roll returns with a much lower turnover compared to the other momentum strategies. An advantage of buying contracts further along the curve is that these can potentially be kept in the portfolio much longer. We observe that applying such a strategy leads to a reduction of more than 50% in turnover and more than doubles the net return to 8.42% per annum compared to a traditional momentum strategy.

Our results are not due to exposure to the commodity market factor or the carry strategy. Also, liquidity seems unlikely to explain the results as even when accounting for liquidity differences through trading costs, reducing the maximum maturity of futures contracts from 12 to 6 months, investing in the most liquid futures contracts, allowing for a one-day implementation lag to reduce trade impact or focusing on the period since 2000, the results remain qualitatively the same.

Chapter 4

Common factors in commodity futures curves

Joint work with Dick van Dijk and Michel van der Wel

4.1 Introduction

Commodities have become a popular asset class among investors in recent years, as they offer interesting diversification opportunities in the context of broader investment portfolios as well as a useful hedge against inflation.[1] While commodity characteristics vary, there are comovements in their prices, as seen in the price boom during the period 2006-2008 and the downturn during the second half of 2008. As commodity prices are important from an economic, sociological, and political perspective, it is important to understand their dynamics and their interrelations. This makes it interesting to look into the possibility to characterize commodity price comovements.

In this paper we examine commonality among commodities in terms of factors driving their entire futures curve, i.e. the collection of all available futures contracts of a particular commodity. By including futures information we can examine not only comovement in the price level of different commodities but also comovement in the shape of their curves. This additional analysis of important futures curve characteristics such as slope and curvature

[1] See e.g. Gorton and Rouwenhorst (2006). Note that the usefulness to act as a hedging instrument varies across individual commodities because of their heterogeneous nature, see Erb and Harvey (2006), Brooks and Prokopczuk (2013), and Daskalaki, Kostakis, and Skiadopoulos (2014), among others.

sheds light on the beliefs of market participants about commodity price fundamentals. For example, the theory of storage states that the slope of the futures curve is related to the level of inventories (Kaldor, 1939; Working, 1949; Fama and French, 1987). Our approach enables us to directly link these slopes across commodities.

We investigate the 24 commodities that constitute the Goldman Sachs Commodity Index (GSCI). These commodities can be split into five sectors: energy, metals, softs, grains, and meats. As a large part of these commodities are also included in the Dow Jones-UBS Commodity Index (DJ-UBSCI),[2] and as we include at each point in time all available futures contracts, our data covers a large part of total exchange-related commodity trading. We use monthly futures prices for the period January 1995 to September 2012.

We use an enhanced version of the Nelson and Siegel (1987) model to extract the factors that drive the commodity futures curves. The Nelson and Siegel (1987) model has good ability to fit the term structure of interest rates (see also Diebold and Li, 2006). As the statistical features of commodity futures curves resemble those of bond yields, the model can be used to characterize commodity futures prices. Although the base model can deal with all kinds of curve shapes we adjust it to be able to handle pronounced seasonal patterns, which are typical for commodities.

To investigate the comovement among commodity futures curves we adopt the framework of Diebold, Li, and Yue (2008) with several extensions. Diebold, Li, and Yue (2008) introduce an additional layer in the Nelson and Siegel (1987) model consisting of unobserved common components that underlie the extracted factors that drive the individual commodity futures curves. We make use of two types of common components, i.e. a market-wide component that is common to all commodities and sector components that are only common to commodities within the same sector. The heterogeneous nature of commodities motivates the inclusion of the sector specific components. Besides the common components there are idiosyncratic components to allow for commodity-specific behavior. In the end we are interested in the relative importance of the common components versus the idiosyncratic components, which is our definition of the degree of commonality.

Our empirical results show that our enhanced version of the Nelson-Siegel model is suited to model commodity futures curves and that there are important common components

[2]This index was formerly known as the Dow Jones-AIG Commodity Index (DJ-AIGCI). Together, the GSCI and DJ-AIGCI are the two commodity indexes that have emerged as industry benchmarks (Stoll and Whaley, 2010).

in the level, slope and curvature factors that characterize these curves. For the commodity level factors, the percentage of explained variation by a combination of the market-wide and sector components is on average 62%. For the slope and curvature factors, these percentages are around 74%. In all cases the commonality is mostly driven by the sector components (between 39% and 50%), which implies that commodities are not an homogeneous asset class. When we investigate commonality over time, we find that common components in our level factors have become increasingly important over our sample period. The percentage of explained variation by common components starts at 53% in 1995 and increases to 68% in 2012. The slope and curvature factors show almost no variation in results over time.

To give a better interpretation to the unobserved components, we investigate their relation with observed economic variables. The market-wide level factor is related to hedging pressure, equity prices and exchange rates. When the net short positions of hedgers increase, the curves' price levels also increase (in line with Gorton et al., 2013). A rise in in equity prices or a weakening of the US dollar also lead to an increase in the common level factor. These findings are in line with, among others, Vansteenkiste (2009), Byrne, Fazio, and Fiess (2012), and Chen, Rogoff, and Rossi (2010). Regarding our market-wide slope component, we find that it increases when the net short positions of hedgers increase. This results in a more backwardated futures curve, which is in line with the theory of normal backwardation (see e.g. Bessembinder, 1992).[3] Furthermore, the slope of the futures curves is negatively related to commodity inventories. This relation between the slope of the futures curve (i.e. convenience yield) and commodity inventories is in line with the theory of storage and findings of Gorton, Hayashi, and Rouwenhorst (2013) and Deaton and Laroque (1992). Thus, using the proposed modeling framework we confirm many findings of existing literature and get additional insights in the commonality of the entire commodity futures curves.

Some papers already use the Nelson-Siegel approach to fit individual commodity futures curves. West (2012) uses it to obtain price estimates for OTC forward contracts beyond the longest available maturity of exchange traded futures contracts. Hansen and Lunde (2013) focus on forecasting and use the Nelson-Siegel framework to predict the term structure of oil futures. An alternative approach to model the commodity futures curve is to express it in terms of unobserved factors and derive futures prices under no-arbitrage conditions (see Gibson and Schwartz, 1990; Schwartz, 1997; Geman, 2005). This approach is more restrictive

[3] Or, e.g., Carter, Rausser, and Schmitz (1983); Chang (1985); de Roon, Nijman, and Veld (2000)

because it requires many assumptions on the market factors and it has difficulty incorporating seasonal patterns (West, 2012). Furthermore, the model complexity makes it difficult to jointly model multiple commodities. Cortazar, Milla, and Severino (2008), Ohana (2010) and Casassus, Liu, and Tang (2013) propose joint models for two commodities but these cannot easily be extended to larger dimensions.[4] As both approaches extract unobserved factors from the futures curve, it is possible that these factors are alike. For all individual commodities, our level and slope factors are similar to their corresponding unobserved spot price and convenience yield factors of the Schwartz (1997) three-factor model, while our curvature factor explains part of his residuals.

There is a large existing literature on comovement of commodity prices. Some work focuses on the drivers of comovement and the macroeconomy. Pindyck and Rotemberg (1990) state in their excess comovement hypothesis that seemingly unrelated commodities comove more than expected, after correcting for macroeconomic influences. Other papers find weaker evidence or reject this excess comovement hypothesis after accounting for model misspecification, conditional heteroskedasticity, and non normality (Deb, Trivedi, and Varangis, 1996) or after incorporating inventory and harvest information (Ai, Chatrath, and Song, 2006). We focus on quantifying the level of commonality across commodities and do not split this commonality in a “normal” and “excess” part.

More closely related to our work is the strand of literature that documents common unobserved factors among individual commodity prices. Vansteenkiste (2009) and Byrne, Fazio, and Fiess (2012) use principal component techniques to extract a latent factor that drives non-fuel commodity prices and link this factor to observed economic variables, like exchange rates or real interest rates. We differ by investigating the comovement of the entire futures curve of commodity prices instead of one price series per commodity. This allows us to investigate comovement, besides levels, in the shapes of futures curves. Furthermore, it allows for a more accurate correction of seasonality effects.

Besides commonality in commodity prices, many papers investigate commonality in commodity returns or their volatilities. Christoffersen, Lunde, and Olesen (2014) document

[4] Cortazar, Milla, and Severino (2008) propose a multi-commodity version of the Cortazar and Naranjo (2006) model for two oil related commodities that uses common and commodity specific factors. Ohana (2010) captures the joint evolution of correlated futures curves by incorporating both the local and global dependence structures between slopes and levels. Casassus, Liu, and Tang (2013) jointly model the convenience yields of two commodities using a multi-commodity feedback affine (MCFA) model to match observed futures correlations.

a clear factor structure in both returns and volatilities. The asset pricing literature focuses on factors that can explain the cross-section of commodity futures returns. Szymanowska, de Roon, Nijman, and van den Goorbergh (2014) show that cross-sectional variation in commodity futures (portfolio) returns can be attributed to a single basis factor for spot premia and to two additional basis factors for term premia. In contrast, Daskalaki, Kostakis, and Skiadopoulos (2014) document that there are no macro, equity or commodity-motivated factors that can price the cross-section of commodity futures.

The rest of this paper is organized as follows. In the next section we present the methodology, followed by Section 4.3 where we present the data and the descriptive statistics. Section 4.44 and 4.5 discuss the empirical results, while Section 4.6 concludes.

4.2 Methodology

In this section we show how the individual commodity futures curves are decomposed into level, slope, curvature, and seasonal factors, and how commonality in these factors across commodities is modeled. The first subsection discusses our model. In the second subsection, we discuss the applied estimation procedure.

4.2.1 Model

We model a collection of futures prices for N different commodities. The futures price with maturity τ (measured in years) for commodity i at time t is denoted by $f_{i,t}(\tau)$. Our starting point is the dynamic version of the Nelson and Siegel (1987) model, as introduced in Diebold and Li (2006), to describe the futures curve of each individual commodity i, for $i = 1, 2, \ldots, N$, as

$$f_{i,t}(\tau) = l_{i,t} + s_{i,t}\left(\frac{1-\exp^{-\lambda_i\tau}}{\lambda_i\tau}\right) + c_{i,t}\left(\frac{1-\exp^{-\lambda_i\tau}}{\lambda_i\tau} - \exp^{-\lambda_i\tau}\right) + \nu_{i,t}(\tau), \qquad (4.1)$$

where $l_{i,t}$, $s_{i,t}$, $c_{i,t}$ are interpreted as time-varying unobserved factors, the decay parameter $\lambda_i > 0$ is assumed to be constant over time, and $\nu_{i,t}(\tau)$ is a disturbance term, where Σ_ν is the covariance matrix of $(\nu_{i,t}(\tau_1), \nu_{i,t}(\tau_2), \ldots, \nu_{i,t}(\tau_{J_i}))'$. The interpretation of the unobserved factors $l_{i,t}$, $s_{i,t}$, and $c_{i,t}$ is determined by their loadings. The loading on the first factor is a constant such that $l_{i,t}$ affects all futures prices in the same way irrespective of their maturity, hence the name level factor. The loading on the second factor is a decreasing function of the futures contract maturity τ and $s_{i,t}$ can therefore be regarded as the slope of the futures curve. The loading on the third factor is a concave function of τ, which allows to fit humped-shaped futures curves, and $c_{i,t}$ can thus be interpreted as a curvature factor.[5]

We make two adjustments to the base model given in (4.1). First, we enhance the model to account for seasonality. Commodity futures curves display pronounced seasonal patterns due to seasonal supply and demand effects, for example related to crop cycles and weather conditions (see e.g. Milonas, 1991). We account for seasonality by including a trigonometric function that depends on the expiry month $g_i(t,\tau)$ of the contract.[6] Trigonometric functions are often used to model seasonality, see for instance Sorensen (2002). We thus enhance the model in (4.1) with a seasonal term given by $\kappa_i \cos(\omega g_i(t,\tau) - \omega\theta_i)$, where the parameter κ_i determines the commodity-specific exposure to the seasonal, the constant ω determines the cycle length, and the parameter θ_i indicates the peak of the seasonal term.

Second, for small values of λ_i, the loadings of the slope factor only slowly decline towards zero as the maturity increases. While this helps to fit futures curves that are very smooth, it becomes difficult to identify the level and slope factors when the term-structure dimension is limited. By re-centering the loadings of the slope and curvature factors to zero at the one year maturity, we make sure that the slope factor does not absorb movements of the curve's level. In this case, the level factor represents the price level of the one year futures contract, while the slope and curvature factor have a similar interpretation as before. The curve is given by

$$\begin{aligned} f_{i,t}(\tau) = & l_{i,t} + s_{i,t}\left[\left(\frac{1-\exp^{-\lambda_i\tau}}{\lambda_i\tau}\right) - \left(\frac{1-\exp^{-\lambda_i}}{\lambda_i}\right)\right] \\ & + c_{i,t}\left[\left(\frac{1-\exp^{-\lambda_i\tau}}{\lambda_i\tau} - \exp^{-\lambda_i\tau}\right) - \left(\frac{1-\exp^{-\lambda_i}}{\lambda_i} - \exp^{-\lambda_i}\right)\right] \\ & + \kappa_i\cos(\omega g_i(t,\tau) - \omega\theta_i) + \nu_{i,t}(\tau). \end{aligned} \tag{4.2}$$

[5]The loading for slope starts at 1 for $\tau \to 0$ and monotonically declines towards zero as the maturity τ increases. The loading for $c_{i,t}$ is equal to 0 for $\tau \to 0$ and $\tau \to \infty$, and reaches a maximum value for maturity τ^*, which depends on the value of λ_i.

[6]The mathematical expression for the expiration month is $g_i(t,\tau) = t + \tau - S\left\lfloor\frac{t+\tau}{S}\right\rfloor$, with $S = 12$ the number of distinct "seasons" and the function $\lfloor x \rfloor$ returns the largest integer not greater than x. Therefore, $g_i(t,\tau)$ results in the integers $\{0, 1, \ldots, 11\}$. Our sample starts in January with $t = 0$ such that the integers $\{0, 1, \ldots, 11\}$ represent the expiry months January, February, ..., December.

The constants we subtract from the slope and curvature loadings are commodity specific (due to λ_i) but fixed over time. The constant only has an effect when λ_i becomes small as for large values of λ_i is close to zero.

We investigate the comovement of commodity prices by linking the futures curves of the individual commodities. This link across commodities is accomplished by decomposing the level, slope, and curvature factors into latent market-wide, sector, and idiosyncratic components. We define the factor decompositions

$$\begin{aligned} l_{i,t} &= \alpha_i^L + \beta_i^L L_{market,t} + \gamma_i^L L_{sector,t} + L_{i,t}, \\ s_{i,t} &= \alpha_i^S + \beta_i^S S_{market,t} + \gamma_i^S S_{sector,t} + S_{i,t}, \\ c_{i,t} &= \alpha_i^C + \beta_i^C C_{market,t} + \gamma_i^C C_{sector,t} + C_{i,t}, \end{aligned} \tag{4.3}$$

where α_i^X for $X = \{L, S, C\}$ are constant terms, β_i^X are loadings on the latent market-wide component $X_{market,t}$, γ_i^X are loadings on sector components $X_{sector,t}$, and $X_{i,t}$ are idiosyncratic components. The (absolute) magnitude of the coefficients β_i^X and γ_i^X determines the degree of comovement across all commodities and across commodities within a specific sector, respectively.

We include sector components besides the market-wide components, because we may expect that commodities in the same sector are more closely related than commodities across different sectors. The three-way decomposition of the futures curve factors is reminiscent of Kose, Otrok, and Whiteman (2003), who let business cycles in different countries depend on global, regional, and an idiosyncratic part, but do not work in the Nelson-Siegel framework. In contrast, Diebold, Li, and Yue (2008) use the Nelson-Siegel framework but decompose the country yield factors only in a global and idiosyncratic component.

The market-wide, sector, and idiosyncratic components are assumed to have first-order autoregressive dynamics

$$\begin{pmatrix} \Delta L_{y,t} \\ S_{y,t} \\ C_{y,t} \end{pmatrix} = \begin{pmatrix} \phi_{11}^y & \phi_{12}^y & \phi_{13}^y \\ \phi_{21}^y & \phi_{22}^y & \phi_{23}^y \\ \phi_{31}^y & \phi_{32}^y & \phi_{33}^y \end{pmatrix} \begin{pmatrix} \Delta L_{y,t-1} \\ S_{y,t-1} \\ C_{y,t-1} \end{pmatrix} + \begin{pmatrix} \eta_{y,t}^L \\ \eta_{y,t}^S \\ \eta_{y,t}^C \end{pmatrix}, \tag{4.4}$$

where $y = \{$*market, sector, idiosyncratic*$\}$, and the disturbances $\eta_{y,t} = (\eta^L_{y,t}, \eta^S_{y,t}, \eta^C_{y,t})$ are normally distributed with covariance matrix Σ_{η_y}. The choice for VAR(1) dynamics for the components is in line with Diebold and Li (2006) and Diebold et al. (2008). One difference is that we assume the level factors to be non-stationary and model them as first differences. Duffee (2011) makes a similar assumption for yield data, while Hansen and Lunde (2013) do the same for oil commodity futures.

To facilitate tractable estimation in our applications and to let all covariance come from common factors, we use a restricted version of the model given by (4.2)-(4.4). That is, we do not allow for cross-correlation of the shocks $\nu_{i,t}(\tau)$ across maturities and commodities, and for the shocks $\eta_{y,t}$ across market-wide, sector, and idiosyncratic components. In other words, all the covariance disturbance matrices Σ_ν and Σ_{η_y} are diagonal. Also the autoregressive matrices in (4.4) are assumed to be diagonal.

As often with factor models, we need to make sure that our unobserved factors are uniquely identified. Here, we have two identification issues as neither the signs nor the scales of the market-wide and sector factors and their factor loadings are separately identified. We follow Sargent and Sims (1977) and Stock and Watson (1989) to identify the scales by assuming that each disturbance variance is equal to a constant, i.e. $\Sigma_{\eta_z}(j,j) = 0.01$ for $z = \{$*market, sector*$\}$ and $j = 1, 2, 3$.[7] We identify the factor signs by restricting one of the loadings for each of the market-wide and sector components to be positive.

4.2.2 Estimation

The model as given by (4.2)-(4.4) can be estimated either using a two-step approach (see, e.g., Diebold et al., 2008) or a one-step approach (see, e.g., Diebold et al., 2006). In the two-step approach one first extracts the latent factors $l_{i,t}$, $s_{i,t}$, and $c_{i,t}$ in (4.2) at each point in time for each commodity and in the second step decomposes the extracted factors into the market-wide, sector and idiosyncratic components. We use the one-step approach where we cast the complete model in a state space representation and use the Kalman filter to estimate all parameters as well as the latent factors and their decomposition simultaneously. The advantage of this procedure is that it takes the estimation uncertainty in the extracted factors into account in their decomposition. Furthermore we can use both time series and

[7] The value of 0.01 is in line with the estimated variance of the idiosyncratic factor' disturbances, as will become clear in the results section.

cross-sectional observations to accurately estimate the parameters. To initialize the one-step approach we can use estimation results of smaller versions of our full model, e.g. a variant without market-wide or sector components.

The state space representation follows naturally from the model given by (4.2)-(4.4). The measurement equation (4.A.1) is given in Appendix 4.A and based on (4.2) and (4.3). The individual latent level $l_{i,t}$, slope $s_{i,t}$, and curvature $c_{i,t}$ factors do not appear in the measurement equation, as we can link the observed futures prices $f_{i,t}(\tau)$ directly to the unobserved market-wide, sector and idiosyncratic components. The transition equations of the latent states are given by (4.4).

Our model can be presented in stacked form if we treat the multivariate series as univariate series, following Koopman and Durbin (2000). We can consider the futures prices separately since we assume that there is no cross-correlation across different commodities (all commonality is absorbed by the factors). The univariate treatment gives us not only computational gains but also allows the number of term-structure observations J_i to be time-varying and deal with our unbalanced panel. We collect all unknown coefficients in the parameter vector Ψ, i.e. the commodity specific parameters in (4.2) and (4.3) $\left(\lambda_i, \kappa_i, \theta_i, \alpha_i^X, \beta_i^X, \gamma_i^X\right)$, and the diagonal elements of the VAR coefficient matrices Φ^y, and the variance matrices Σ_ν and Σ_{η_y}. Estimation of Ψ is based on the numerical maximization of the loglikelihood function that is constructed via the prediction error decomposition.

4.3 Data

We study futures curves for 24 commodities. We consider the period January 1995 to September 2012 and use all individual futures contracts that expire between January 1995 and December 2030.[8] Our commodity selection is based on the composition of the S&P Goldman Sachs Commodity Index (GSCI). The GSCI constituents can be split in five sectors: energy, metals, softs, grains, and meats. An overview of the data (as obtained from Thomson Reuters Datastream) is given in Table 4.1. All our analyses are done at the monthly frequency, and for this we use month-end log settlement prices. All prices are standardized

[8]The start of the sample period is based on the availability of the metal commodities traded on LME, which are only available from July 1993 onward.

since the pricing grid of the commodities is quite diverse.[9] To avoid liquidity issues we do not consider price information of (i) contracts with a monthly return that equals zero, and (ii) contracts in the expiration month. Furthermore, we filter out data errors by excluding contracts that have abnormal returns compared to adjacent contracts. These filters lead to exclusion of approximately 1.3% of the data.

Table 4.1 Commodity data overview

The table presents an overview of the 24 commodity futures series that are all present in the S&P Goldman Sachs Commodity Index (GSCI). The sector classification is the same as GSCI. We consider the period January 1995 to September 2012. We show the number of cross-sectional contracts that are available in the first and last year of our dataset. Furthermore we show the average annualized return and its volatility of both the first maturing futures contract and the futures contract with 12 months to maturity.

		# contracts		1st nearby contract		12M contract	
Sector	Commodity	Begin	End	$\bar{r}$	$\sigma(r)$	$\bar{r}$	$\sigma(r)$
Energy	Brent crude oil	12	56	12.9%	32.3%	13.4%	19.3%
	WTI crude oil	26	67	9.0%	32.8%	7.8%	17.1%
	Gasoil	14	30	12.8%	32.8%	7.3%	20.7%
	Heating oil	17	19	10.7%	34.8%	6.6%	25.2%
	Natural gas	20	83	-20.6%	53.1%	-1.4%	17.4%
	Gasoline	10	24	18.8%	39.2%	7.6%	23.3%
Metals	Gold	19	19	5.9%	16.4%	4.8%	15.9%
	Silver	19	19	8.4%	30.6%	8.5%	29.8%
	Aluminum	11	60	-4.9%	19.9%	7.8%	17.1%
	Copper	11	60	7.5%	27.6%	20.5%	31.4%
	Lead	11	29	4.7%	30.2%	10.8%	35.1%
	Nickel	11	30	4.2%	36.4%	10.3%	33.7%
	Zinc	11	30	-2.2%	27.6%	14.2%	29.7%
Softs	Cocoa	8	10	0.6%	31.4%	-4.6%	24.9%
	Coffee	7	10	-5.3%	36.7%	-5.6%	28.9%
	Cotton	9	11	-6.9%	29.9%	-2.0%	18.1%
	Sugar	7	8	0.5%	38.2%	8.6%	18.6%
Grains	Corn	8	14	-4.4%	29.0%	6.6%	12.6%
	Soybeans	11	17	7.3%	26.6%	3.1%	17.8%
	Chicago wheat	7	10	-10.2%	30.3%	4.4%	17.5%
	Kansas wheat	5	9	-0.2%	29.3%	1.3%	21.5%
Meats	Feeder cattle	8	8	1.5%	14.5%	4.2%	10.3%
	Lean hogs	8	9	-7.2%	28.0%	7.4%	14.1%
	Live cattle	7	7	0.5%	14.9%	3.8%	8.5%

Table 4.1 shows that the number of available contracts (term-structure observations) varies per commodity. This is caused by differences in (i) the number of distinct expiration months a year or (ii) the maximum time-to-maturity. Energy and industrial metal com-

[9] We standardize prices by setting the first nearby contract price in January 1995 of all commodities equal to 100. All other prices are adjusted such that (time series and term structure) returns remain unchanged.

modities have an expiring futures contract each month and also contracts with long dated maturities. In contrast, agricultural commodities have a small number of active futures because they have only five to eight distinct expiry months a year. Furthermore, the maximum maturity of these contracts is between one and two years, which is in line with the length of their crop cycles and storability of these commodities. The variation in the number of contracts and the maximum time to maturity indicate that it is important to use a commodity specific decay parameter λ_i in (4.2). Besides differences in the number of available contracts across commodities, we notice the same within commodities over time. Especially the number of contracts for energy and industrial metal commodities greatly increases over our 17 years sample. Even though our estimation methods can deal with this increase of available contracts over time, our choice to use a fixed λ parameter limits the model flexibility. Instead of allowing λ to vary over time (see e.g. Koopman, Mallee, and Van der Wel, 2010), we choose to introduce a maturity bound to exclude long-dated contracts. By limiting the term-structure dimension variation within a commodity, we can keep using a fixed λ value. Furthermore, these long-dated contracts are possibly less liquid and hence have more noisy price information, which could otherwise affect our results. The introduced maturity bound excludes on average 10% of our observations.[10]

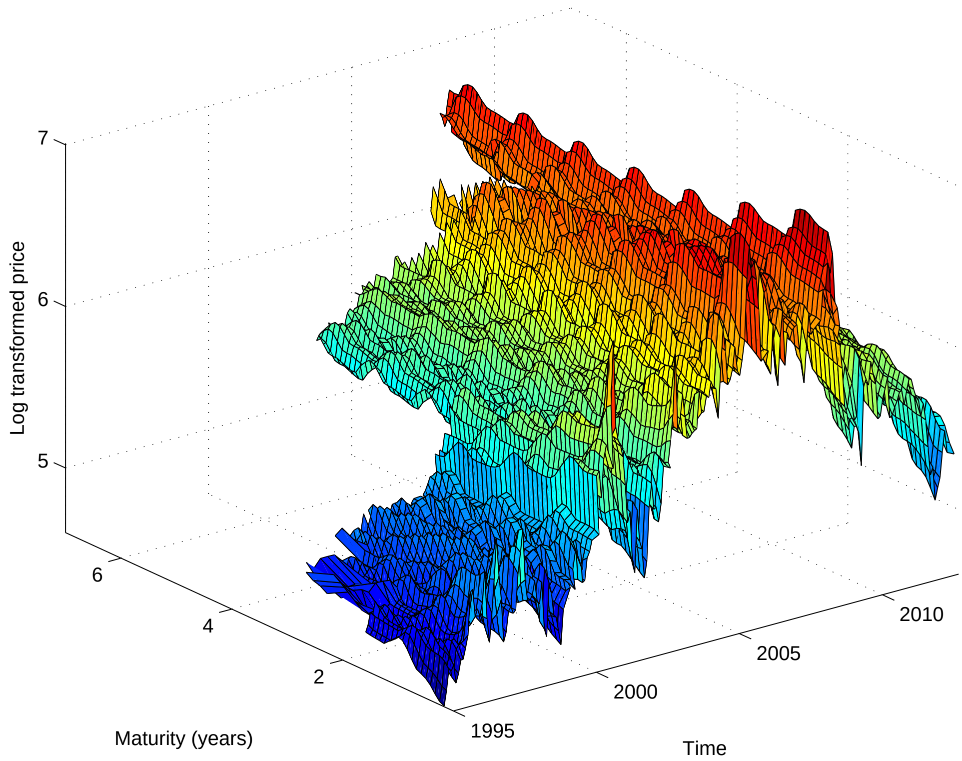
The summary statistics in Table 4.1 show that there are large return differences both across commodities and along their futures curves. The returns of the contracts range between -20.6% to 20.5% and are more extreme for the first nearby contract. The volatility of the returns confirms this as in almost all cases the 12 month contract returns are less volatile than the returns of the first nearby contract, also known as the Samuelson (1965) effect.

[10] Appendix 4.B provides additional details on the introduced maturity bound.

Figure 4.1 Commodity futures curves

These figures gives insight in the data by showing the complete set of available futures prices for natural gas and coffee. The figures show the commodity futures curves at each month in time.

(a) Natural gas



(b) Coffee

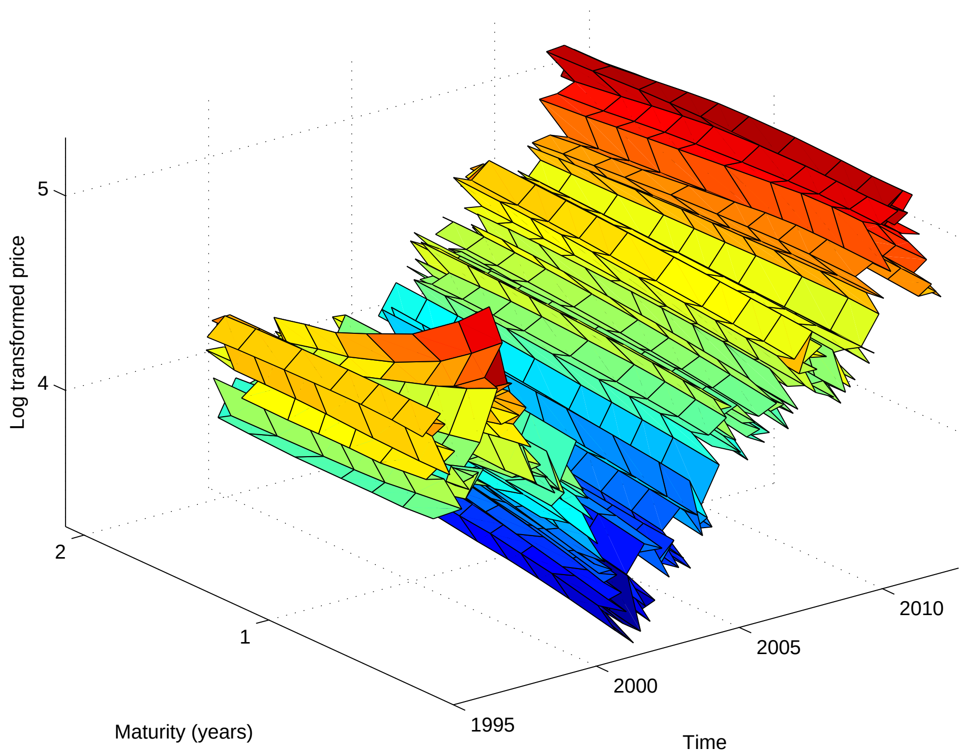


Figure 4.1 gives some insight in the data we use by showing the complete set of available futures prices for natural gas and coffee. For both commodities we observe that the shape of the futures curve varies substantially over time, with alternating periods of pronounced contango and backwardation especially for natural gas.[11] Both futures curves also clearly show the large increase in the general price level during the period 2005-2008. A notable difference between these commodities is that the futures curve of natural gas displays a strong periodic pattern with spikes occurring for expiry months during the winter, while the curve of coffee does not show any signs of seasonality. Finally, Figure 4.1 illustrates that the number of available contracts varies over time. For both natural gas and coffee (and in fact also for most other commodities), contracts with longer maturities only have become available in the most recent years of our sample period.

4.4 Individual commodity results

Before we estimate the full market-wide state space model, we start with analyzing the individual commodities separately. We apply our Nelson-Siegel set-up to each individual commodity, i.e. we leave out the market-wide and sector components in (4.3). We decide on the exact model specification to be used for each commodity based on the parameter estimates and the extracted factors. Furthermore, we show that our framework is suitable for modeling commodity futures data and compare our unobserved level and slope factors with the latent spot price and convenience yield in the Schwartz (1997) three-factor model.

4.4.1 Estimation results individual factors

In Section 4.2, we presented the model in general form, where all commodity curves are built up from a level, slope, and curvature factor combined with a seasonal term. However, not all commodity curves may show dynamics for which the flexibility of three factors is needed. Furthermore, not necessarily all commodities display periodic behavior. Based on the features of each specific commodity we decide on the number of factors to include and whether to include the seasonal term or not.

[11] An upward sloping commodity futures curve is said to be in contango, while a downward sloping curve is in backwardation.

Table 4.2 Individual commodity state space results

The table presents the estimation results and the fit of the individual commodity state space models. We show the final specification we use for each individual commodity, whereby 2F stands for a model with only a level and slope factor, 3F also includes a curvature factor, and 3FS adds a seasonal correction term. The model fit is shown in terms of R^2. The estimated parameter values for the decay parameter λ, the exposure κ to the seasonal term, and the most expensive contract expiry month θ, where $\theta = 0$ corresponds to January. The last column presents the volatility of errors $\sigma(\nu)$. Standard errors of all estimates are provided between brackets. For models that do not contain a seasonal component, the κ and θ parameters are irrelevant, which is represented by a horizontal dash.

Sector	Commodity	Model	R^2	λ		κ		θ		$\sigma(\nu)$
Energy	Brent crude oil	3F	99.99%	1.144	(0.01)	-	-	-	-	0.65%
	WTI crude oil	3F	99.99%	1.272	(0.01)	-	-	-	-	0.79%
	Gasoil	3FS	99.98%	1.688	(0.04)	0.92	(0.02)	0.4	(0.00)	1.03%
	Heating oil	3FS	99.97%	3.600	(0.03)	2.80	(0.01)	0.8	(0.00)	1.21%
	Natural gas	3FS	99.40%	1.137	(0.02)	6.38	(0.01)	0.9	(0.00)	3.56%
	Gasoline	3FS	99.94%	3.285	(0.04)	4.78	(0.01)	6.2	(0.00)	1.52%
Metals	Gold	2F	99.99%	0.011	(6.36)	-	-	-	-	0.44%
	Silver	2F	99.99%	0.026	(1.79)	-	-	-	-	0.54%
	Aluminum	2F	99.91%	0.187	(0.04)	-	-	-	-	0.77%
	Copper	2F	99.97%	0.111	(0.10)	-	-	-	-	1.13%
	Lead	2F	99.99%	0.324	(0.13)	-	-	-	-	0.57%
	Nickel	2F	99.98%	0.095	(0.28)	-	-	-	-	0.91%
	Zinc	2F	99.97%	0.059	(0.19)	-	-	-	-	0.69%
Softs	Cocoa	3F	99.97%	1.413	(0.03)	-	-	-	-	0.58%
	Coffee	3F	99.98%	1.468	(0.03)	-	-	-	-	0.57%
	Cotton	3FS	99.74%	3.482	(0.03)	0.91	(0.06)	5.9	(0.01)	1.21%
	Sugar	3FS	99.91%	3.272	(0.04)	1.13	(0.05)	2.5	(0.01)	1.32%
Grains	Corn	3FS	99.81%	2.743	(0.04)	1.80	(0.03)	5.7	(0.00)	1.56%
	Soybeans	3FS	99.89%	3.350	(0.03)	1.30	(0.03)	5.7	(0.00)	1.19%
	Chicago wheat	3FS	99.84%	1.496	(0.17)	1.53	(0.04)	2.0	(0.00)	1.51%
	Kansas wheat	3FS	99.86%	2.461	(0.08)	1.44	(0.05)	2.2	(0.01)	1.35%
Meats	Feeder cattle	3FS	99.91%	5.221	(0.07)	0.73	(0.07)	10.0	(0.01)	0.71%
	Lean hogs	3FS	97.80%	4.109	(0.07)	7.78	(0.02)	6.1	(0.00)	3.09%
	Live cattle	3FS	99.60%	4.487	(0.10)	2.32	(0.03)	1.9	(0.00)	1.29%

Figure 4.2 Model fit commodity futures curves

These figures show example of the fit of our individual models. The crosses represent the observed price data while the lines correspond to the fitted values of our models. The raw prices are first standardized and thereafter we apply a log-transformation. The horizontal axis shows the time to maturity (τ) in years.

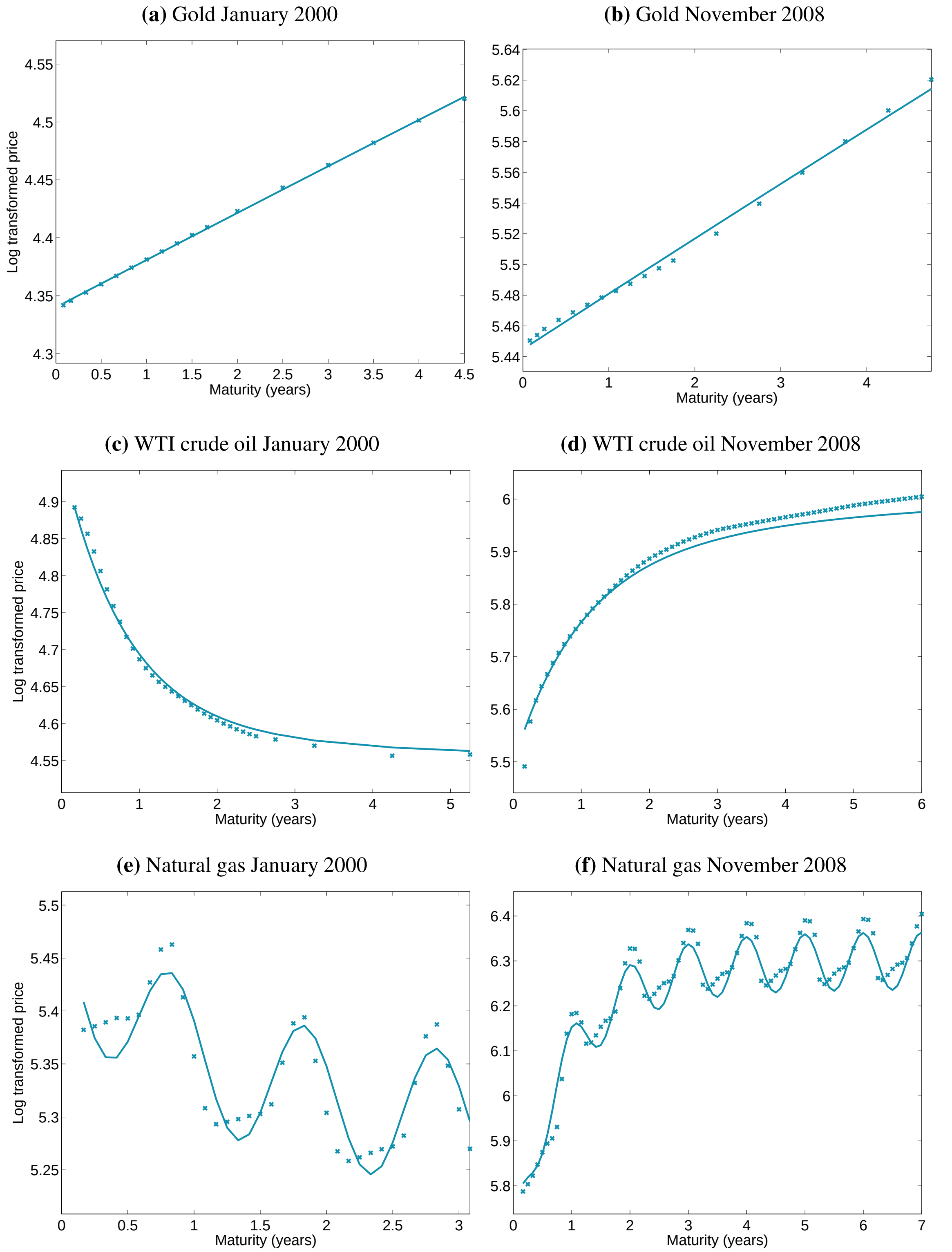


Table 4.2 presents the final model choice and the estimated parameter values for each of the commodities. The final model choice is given in column 3 and is based on several criteria.[12] First, we compare the results of three factor models with and without seasonal term, to see if the exposure κ to the seasonal correction is significantly different from zero. The 13 commodities with the label "3FS" have highly significant κ parameters, ranging between 0.73 and 7.78. Most of these commodities have a clear crop cycle, i.e. a seasonal supply, while others, like natural gas, have well-known seasonal demand. We include three factors for all commodities with periodic behavior because even after the seasonal correction their curves display a large variety of shapes. Second, we need to decide if we include a curvature factor for the remaining 11 commodities. Based on Akaike Information Criterion or Bayesian Information Criterion values, we should choose to include all three factors. However, low λ values for metal commodities lead to slope and curvature loadings that are highly correlated, which lead to identification problems. Hence, we decide to exclude the curvature factor for the metal commodities.

The fourth column in Table 4.2 shows the in-sample fit of our models. With the exception of lean hogs all R^2 values are above 99.7%, which supports the use of the Nelson-Siegel framework for commodity futures prices. Figure 4.2 shows representative examples of the model fit for three of the 24 commodities. Subfigures A and B correspond to the gold futures curve, Subfigures C and D correspond to the curve of WTI crude oil, while Subfigures E and F correspond to natural gas data. The crosses represent the observed price data while the lines correspond to the fitted values of our models. Note that the presented figures are snapshots at one particular point in time. For each of these three commodities, we use a different version of our model. The futures curves of gold are almost straight lines. Hence, we can easily fit the prices with only a level and slope factor and we do not need a curvature factor or a seasonal term. The futures curves of WTI crude oil do not show seasonal patterns but do have a curved shape. In a static case, it would be possible to fit this curve with just a level and a slope factor. However, curves change over time and a two factor model is not flexible enough to cope with these changes. Below in Figure 4.4 and 4.5 we show that a three factor model is more suitable. The natural gas futures curve displays a pronounced seasonal pattern. Therefore, we use all three factors (level, slope and curvature) plus a seasonal term. In general our fitted values are close to the real prices, with some exceptions at the short end

[12] Appendix 4.C provides more detailed results.

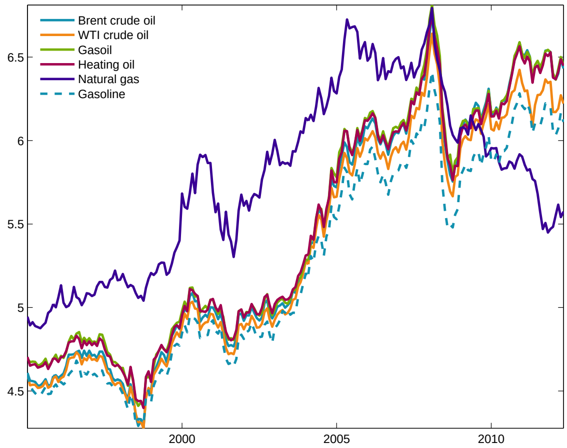
of the curve. The inclusion of the seasonal term seems an appropriate solution to model the periodic behavior.

The remaining columns of Table 4.2 show the parameter estimates with corresponding standard errors. The decay parameter λ determines the shape of the slope and curvature loadings. Large values of λ lead to quickly declining slope loadings and move the peak of the curvature loadings to the short-end of the curve. When $\lambda = 2$, the slope loading starts at 0.49, it declines to zero for $\tau = 1$ (by construction), and is equal to -0.19 and -0.33, for $\tau = 2$ and $\tau = 5$ years, respectively. The curvature loading peaks at 11 months maturity. By contrast, when lambda is equal to 0.01 the slope loading starts at 0.005, it gradually declines towards zero for $\tau = 1$, and is equal to -0.005 and -0.020, for $\tau = 2$ and $\tau = 5$ years, respectively. The curvature loadings are almost a mirror image of the slope loadings, which is why we exclude the third factor for the metal commodities. The maximum curvature loading (when lambda is equal to 0.01) is achieved for $\tau = 179.33$, i.e. 179 years and 4 months, which is way beyond the highest maturity that is included in our sample. The decay parameter λ varies substantially across commodities, ranging from 0.011 for gold to 5.2 for feeder cattle. A value of 0.010 is the lowest value we allow to prevent that the loadings of the level and slope factor become too similar. The variation in λ is both due to differences in curve shapes and maximum contract maturity. The effect of different futures' curves shapes on λ becomes clear when we compare results for commodities with a similar maximum contract maturity. Gasoil and soybeans both have futures up to 2.5 years until maturity, while their λ values differ greatly (1.688 versus 3.350, respectively). Related to the seasonal correction term are the exposure κ and the location parameter θ. The θ estimates imply that gasoil, heating oil and natural gas contracts are most expensive between January and February, while most agricultural commodities are more expensive two months before their harvest. The last column shows the volatility of the errors. The volatilities of the errors are small, especially compared to the highly volatile observed prices.

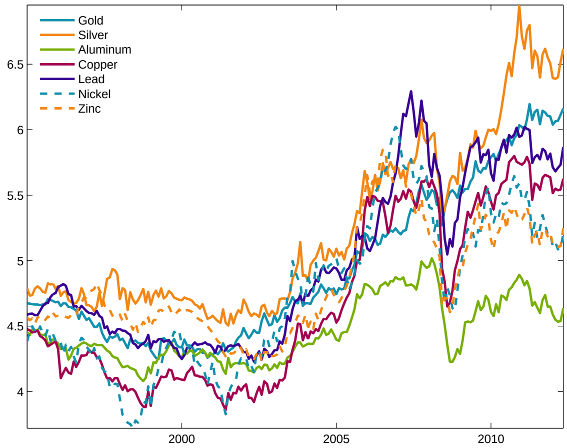
Figure 4.3 Individual models - extracted commodity level factors

These figures show the extracted level factors based on individual models applied to all 24 commodities. Each subfigure shows the estimated level factors for the commodities of a specific sector.

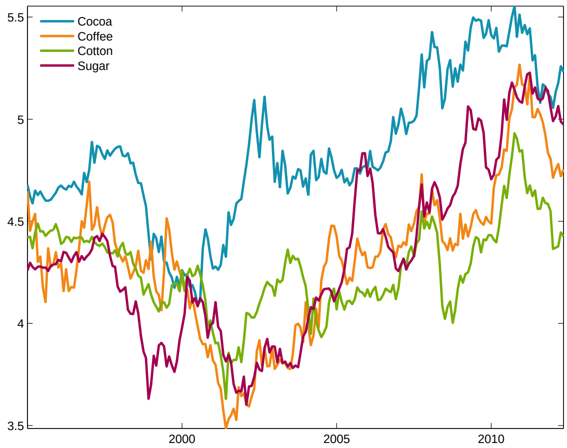
(a) Level - energy



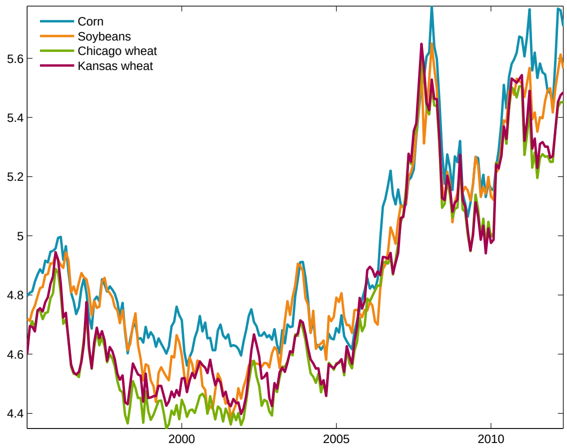
(b) Level - metals



(c) Level - softs



(d) Level - grains



(e) Level - meats

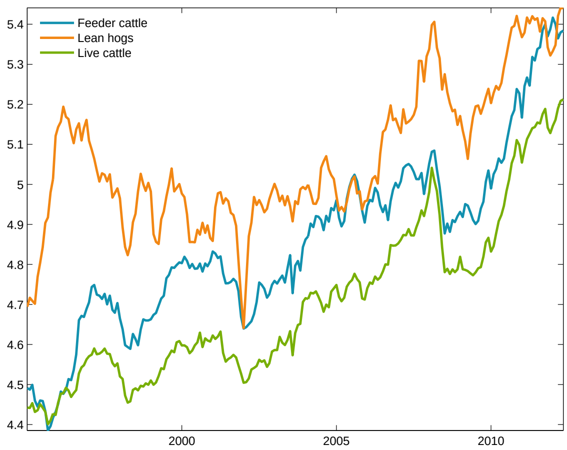
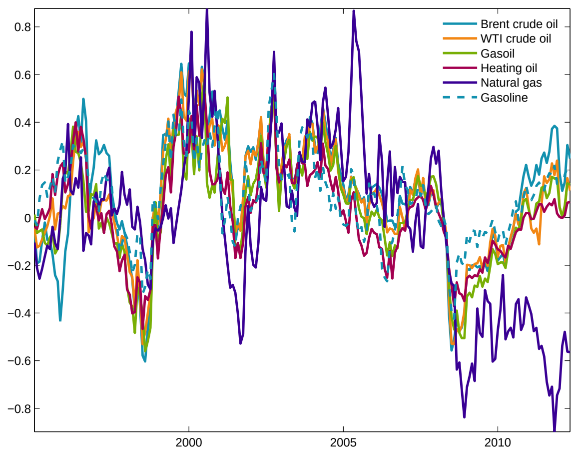


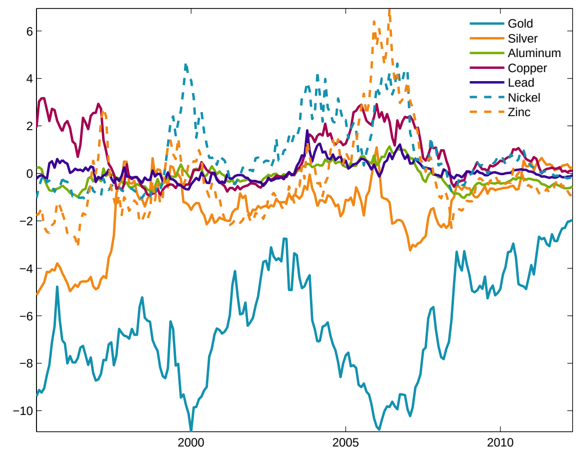
Figure 4.4 Individual models - extracted commodity slope factors

These figures show the extracted slope factors based on individual models applied to all 24 commodities. Each subfigure shows the estimated slope factors for the commodities of a specific sector.

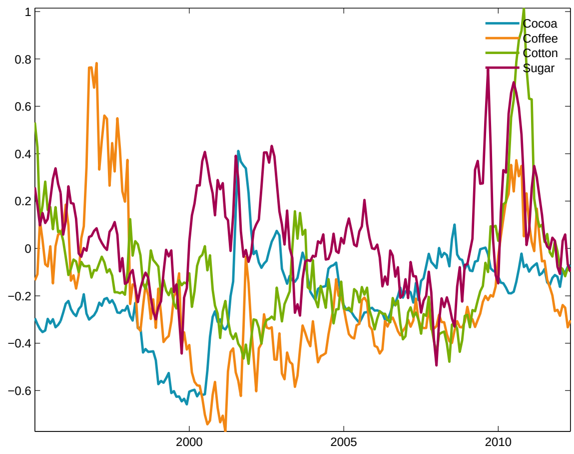
(a) Slope - energy



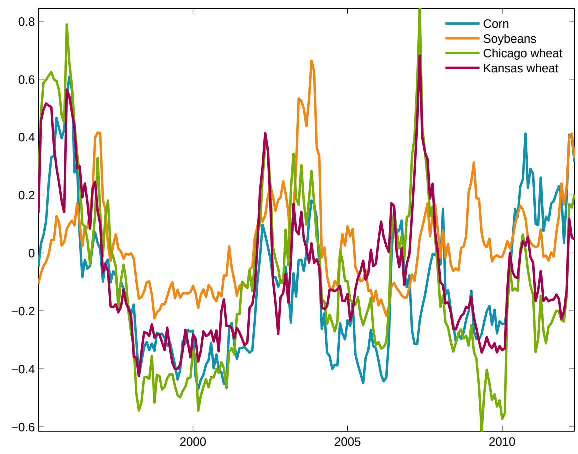
(b) Slope - metals



(c) Slope - softs



(d) Slope - grains



(e) Slope - meats

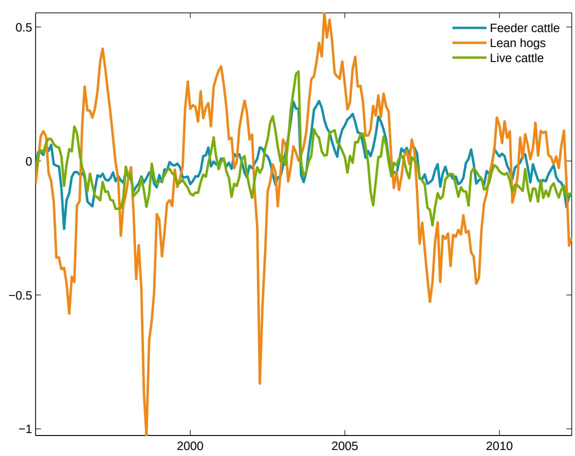
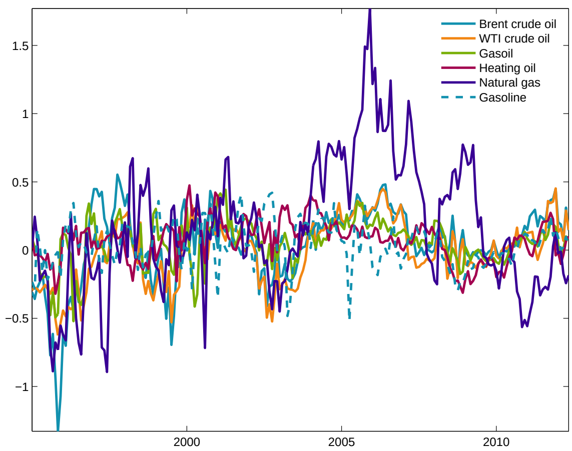
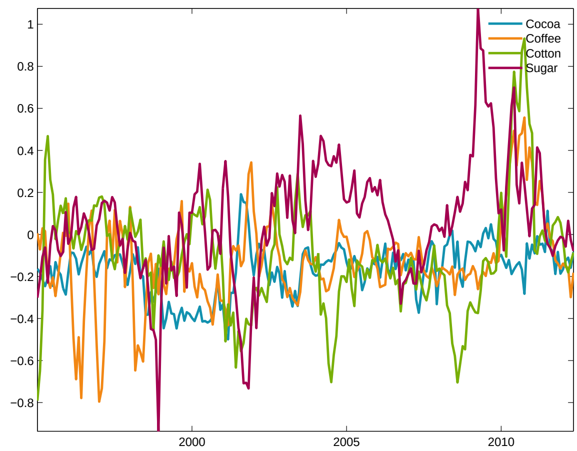


Figure 4.5 Individual models - extracted commodity curvature factors
These figures show the extracted curvature factors based on individual models applied to all 24 commodities. Each subfigure shows the estimated curvature factors for the commodities of a specific sector.

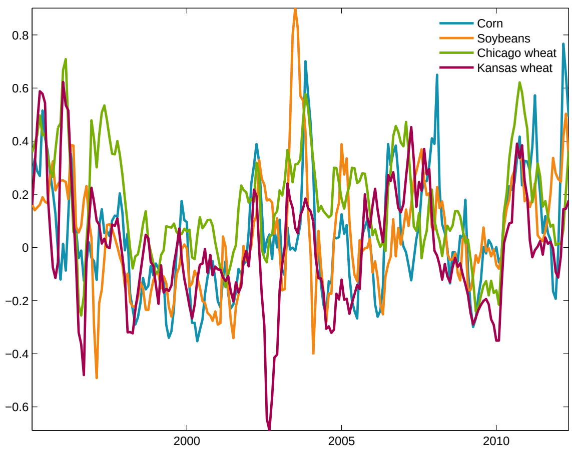
(a) Curvature - energy



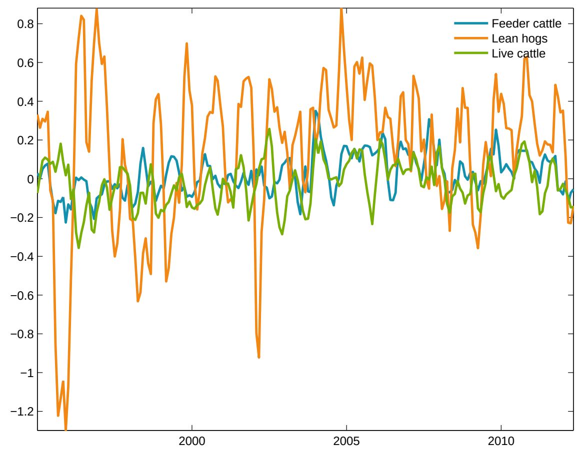
(b) Curvature - softs



(c) Curvature - grains



(d) Curvature - meats



Figures 4.3, 4.4, and 4.5 show the extracted level, slope, and curvature factors per commodity sector. In general, we see a similar pattern in all level factors. Until 2004 they are relatively constant, then they increase until they peak in 2008, whereafter they again remain constant. The level factors within the energy, metals and grains sectors seem to comove the most. The slope factors in Figure 4.4 show some peaks and troughs. Especially for the energy commodities we see a sharp decline in 2008 and a gradual increase thereafter. As the Nelson-Siegel loading on the slope factor in (4.2) is a decreasing function of maturity, a negative factor estimate signifies an upward sloping (i.e. contangoed) futures curve. This implies that in 2008 all the backwardated energy futures curves quickly went into contango,

and only gradually returned back to being backwardated. Last, the curvature factors in Figure 4.5 show again some degree of comovement. Note that the metal commodities are missing, as we find that two factors are enough to capture their curve dynamics. Of the four sectors the energy commodities have the strongest comoving curvature factors. Based on the plots of the individual factors, there seems to be commonality across commodities. The advantage of our framework is that we can easily accommodate this.[13]

4.4.2 Comparison with the Schwartz (1997) three-factor model

Existing models like in Gibson and Schwartz (1990), Schwartz (1997) and Schwartz and Smith (2000) all begin by assuming a functional form for a set of underlying state variables. The futures curve can be derived from these state variables under no arbitrage conditions. Even though our framework is different, both approaches assume that commodity prices are driven by unobserved factors. Hence it is possible that the extracted unobserved factors from both methods are similar, e.g. through factor rotation. In this section we compare the results of our Nelson-Siegel type model with the three-factor model described in Schwartz (1997).

Extending the model of Gibson and Schwartz (1990), Schwartz (1997) assumes that commodity prices are driven by three stochastic factors namely the commodity spot price, the convenience yield and the interest rate.[14] Variations on this approach are given by many subsequent papers on commodity prices (see among others Schwartz and Smith, 2000; Casassus and Collin-Dufresne, 2005). Both the log spot price and the convenience yield are assumed to be mean reverting. Also the instantaneous interest rate is assumed to follow a mean reverting process as in Vasicek (1977).

Schwartz (1997) estimates a simplified version of his three factor model by assuming that the interest rate process is independent of both commodity processes. He first estimates the interest rate parameters and then plugs these into the model.[15] The loadings on the

[13]As a preliminary approach we apply Principal Component Analysis (PCA) both on all commodities and on subgroups that correspond to the commodity sectors. We find commonality across all commodities and within subgroups. Detailed results are shown in Appendix 4.C

[14]Brennan (1991) defines the convenience yield as “the flow of services which accrues to the owner of a physical inventory but not to the owner of a contract for future delivery”

[15]For the complete model specification we refer to page 933 in Schwartz (1997). The interest rate process is estimated separately from the spot price and convenience yield processes, and is based on an observed 3-month Treasury Bill series. Essentially, this simplified version contains only two unobserved states, while still allowing for a time-varying interest rate. When estimating his model we follow the same estimation methodology.

unobserved log spot price and the instantaneous convenience yield show great resemblance with the loadings on our level and slope factors.

Figure 4.6 shows the unobserved level factor of our model and the unobserved spot price of the three-factor Schwartz (1997) model, while Figure 4.7 compares our slope factor with the unobserved convenience yield. Both models are estimated using the same dataset. The similarities are very clear in both figures. The level factor and the spot price both show, in general, an increasing trend with a pronounced dip around the recent financial crisis. The slope and convenience yield factors show more peaks and troughs, which implies upward and downward sloping futures curves. The resemblance of all lines is confirmed by the pair-wise correlations. The average (median) correlation is 0.67 (0.79) between the level factor and the spot price, and 0.76 (0.86) between the slope factor and the convenience yield, respectively. Concluding, our statistical factors level and slope are strongly related to the spot price and convenience yield.

Although our level and slope factors show great resemblance with Schwartz' factors, we have an additional third factor, namely curvature. This gives us additional flexibility to better fit the observed futures prices. In terms of R^2 we increase the model fit by 5.6%.[16] When we examine the residuals of the Schwartz (1997) model, we find that for most commodities there is a strong common factor present. These common factors have on average a correlation of 0.20 with our corresponding curvature factors.

[16] Additional comparison results are presented in Appendix 4.D.

Figure 4.6 Comparison level factor and spot price

These figures show the unobserved level factor of our Nelson-Siegel type models and the unobserved spot price series of the three-factor Schwartz (1997) model. The blue line is the unobserved spot price series and the orange line is our level factor.

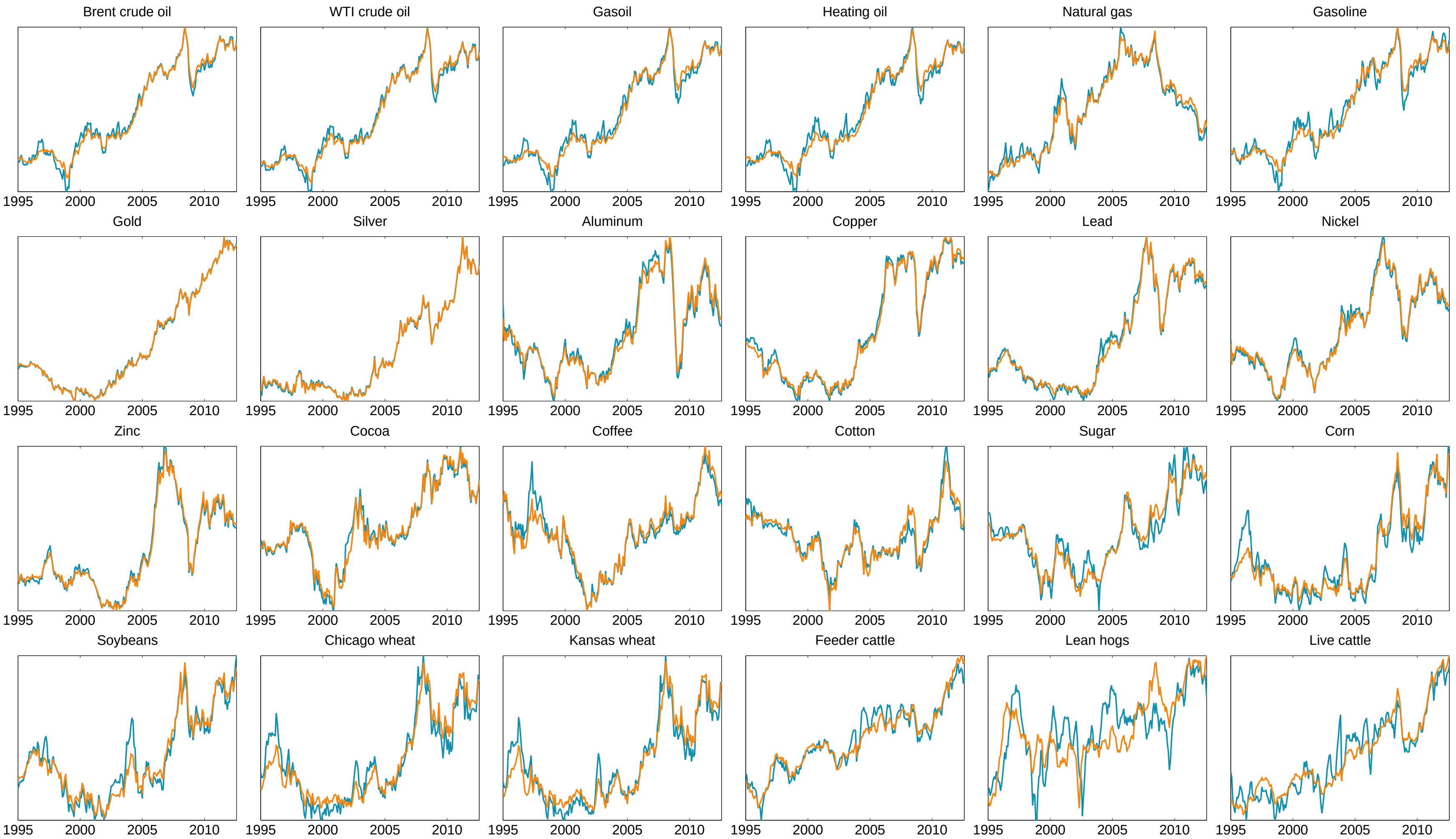
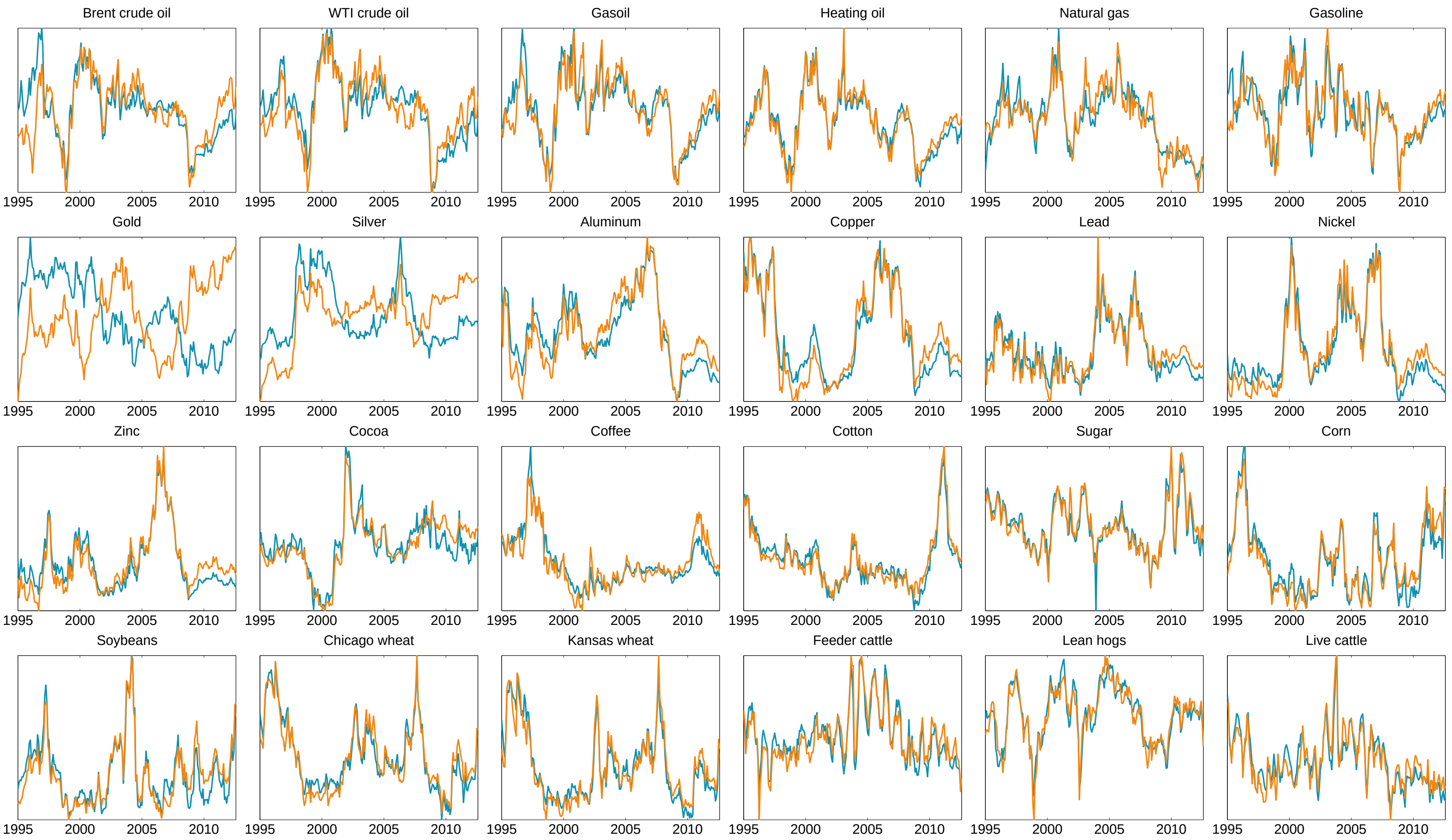


Figure 4.7 Comparison slope factor and convenience yield

These figures show the unobserved slope factor of our Nelson-Siegel type model and the unobserved convenience yield series of the three-factor Schwartz (1997) model. The blue line is the unobserved convenience yield and the orange line is our slope factor.



4.5 Joint model for commodity curves

In this section we return to the market-wide state space model given by (4.2)-(4.4). First, we discuss the commonality across commodities and investigate the importance of the various components using variance decompositions. Thereafter, we investigate the unobserved factors' dynamics and give an interpretation to the unobserved common components. We end this section with the examination of possible variation in the degree of commonality over time and the importance of term structure information.

4.5.1 Commonality results

When we estimate our full model, we fix the λ, κ, and θ parameters to their estimates based on commodity specific data. This reduces the computational burden as parts of the measurement equation are now constant. All other parameters are estimated using the Kalman filter and maximum likelihood.

Table 4.3 Joint model factor loadings

The table presents the estimated loadings on various components in our joint model. Each level, slope, and curvature factor is decomposed in a constant part α, a market-wide part with loading β, a sector part with loading γ, and a commodity specific part. Standard errors of all estimates are provided between brackets. The commodity specific variance estimates, $\sigma^2_{\nu_i}$, are multiplied by 1,000 for readability reasons. Note that we do not include a curvature factor for the metal commodities, which is represented by a horizontal dash.

	Level						Slope						Curvature						Variance	
Commodity	α		β		γ		α		β		γ		α		β		γ		$\sigma^2_{\nu_i}$	
Brent crude oil	4.59	(0.06)	0.81	(0.09)	1.79	(0.09)	0.10	(0.08)	2.99	(0.03)	−0.90	(0.43)	0.06	(0.07)	0.11	(4.23)	3.88	(0.20)	0.01	(0.55)
WTI crude oil	4.58	(0.06)	0.78	(0.15)	1.79	(0.10)	0.08	(0.08)	2.92	(0.18)	−0.70	(0.43)	−0.02	(0.06)	0.36	(0.79)	3.26	(0.23)	0.02	(0.34)
Gasoil	4.70	(0.06)	0.78	(0.15)	1.82	(0.09)	0.03	(0.08)	2.83	(0.17)	−0.10	(0.42)	0.05	(0.04)	1.32	(0.53)	1.81	(0.29)	0.06	(0.35)
Heating oil	4.69	(0.06)	0.75	(0.15)	1.78	(0.09)	0.02	(0.07)	2.44	(0.21)	1.25	(0.38)	0.07	(0.05)	1.34	(0.54)	−0.54	(0.34)	0.12	(0.24)
Natural gas	4.84	(0.07)	0.46	(0.16)	0.81	(0.09)	−0.06	(0.11)	1.51	(0.39)	1.43	(0.15)	0.19	(0.13)	0.83	(0.86)	0.54	(0.64)	1.15	(0.04)
Gasoline	4.57	(0.06)	0.79	(0.15)	1.79	(0.10)	0.08	(0.06)	2.10	(0.18)	0.51	(0.36)	0.05	(0.03)	−0.95	(0.60)	0.12	(0.40)	0.17	(0.23)
Gold	4.66	(0.05)	0.50	(0.10)	0.40	(0.10)	−6.78	(1.67)	−3.09	(1.34)	−2.51	(1.61)	-	-	-	-	-	-	0.02	(0.59)
Silver	4.71	(0.08)	1.12	(0.17)	0.90	(0.16)	−2.12	(1.51)	−1.60	(0.83)	1.30	(0.98)	-	-	-	-	-	-	0.02	(0.53)
Aluminum	4.60	(0.05)	0.49	(0.11)	1.06	(0.04)	−0.15	(0.19)	0.30	(0.33)	3.26	(0.05)	-	-	-	-	-	-	0.03	(0.29)
Copper	4.50	(0.07)	0.78	(0.16)	1.81	(0.13)	0.60	(0.40)	0.47	(0.50)	3.77	(0.57)	-	-	-	-	-	-	0.09	(0.17)
Lead	4.63	(0.08)	0.60	(0.17)	1.41	(0.15)	0.05	(0.20)	−0.86	(0.53)	3.08	(0.55)	-	-	-	-	-	-	0.02	(0.65)
Nickel	4.62	(0.09)	0.69	(0.21)	1.66	(0.19)	0.65	(0.55)	2.26	(1.18)	8.71	(1.22)	-	-	-	-	-	-	0.04	(0.36)
Zinc	4.65	(0.07)	0.65	(0.16)	1.67	(0.13)	−0.34	(0.86)	−0.34	(1.34)	13.57	(1.49)	-	-	-	-	-	-	0.04	(0.35)
Cocoa	4.67	(0.08)	0.78	(0.16)	0.37	(0.25)	−0.20	(0.08)	−0.17	(0.12)	0.17	(0.13)	−0.17	(0.03)	−0.17	(0.33)	−0.29	(0.37)	0.02	(1.07)
Coffee	4.64	(0.08)	0.85	(0.18)	0.70	(0.21)	−0.19	(0.11)	−0.46	(0.25)	3.43	(0.03)	−0.11	(0.06)	−0.58	(0.53)	0.83	(0.56)	0.01	(1.34)
Cotton	4.42	(0.06)	1.02	(0.13)	0.54	(0.25)	−0.06	(0.13)	−0.10	(0.21)	−0.31	(0.22)	−0.10	(0.08)	−0.67	(0.58)	4.29	(4.54)	0.17	(0.31)
Sugar	4.50	(0.07)	0.61	(0.15)	0.36	(0.21)	0.05	(0.07)	−0.13	(0.27)	0.20	(0.27)	0.06	(0.08)	−1.54	(0.64)	0.23	(0.79)	0.17	(0.37)
Corn	4.70	(0.06)	1.65	(0.11)	0.21	(0.26)	−0.10	(0.08)	0.04	(0.21)	1.60	(0.07)	0.06	(0.05)	−0.57	(0.72)	2.79	(0.08)	0.26	(0.24)
Soybeans	4.67	(0.06)	1.48	(0.11)	0.14	(0.12)	0.03	(0.06)	0.05	(0.16)	0.90	(0.17)	0.05	(0.05)	−1.93	(0.66)	2.49	(0.56)	0.15	(0.29)
Chicago wheat	4.53	(0.06)	1.34	(0.13)	1.42	(0.10)	−0.07	(0.12)	0.08	(0.27)	3.72	(0.22)	0.16	(0.05)	2.08	(0.71)	3.04	(0.52)	0.28	(0.27)
Kansas wheat	4.52	(0.06)	1.32	(0.13)	1.42	(0.10)	−0.05	(0.09)	0.10	(0.21)	2.61	(0.18)	−0.00	(0.05)	2.78	(0.80)	3.73	(0.58)	0.23	(0.32)
Feeder cattle	4.58	(0.03)	−0.02	(0.07)	0.86	(0.03)	−0.01	(0.02)	−0.02	(0.11)	1.26	(0.09)	0.02	(0.02)	0.93	(0.33)	0.87	(0.17)	0.03	(0.93)
Lean hogs	4.71	(0.04)	0.43	(0.10)	0.34	(0.09)	−0.04	(0.07)	0.69	(0.33)	0.86	(0.38)	0.14	(0.11)	0.55	(1.28)	−1.68	(1.17)	1.09	(0.13)
Live cattle	4.48	(0.02)	0.34	(0.05)	0.61	(0.04)	−0.03	(0.03)	−0.06	(0.15)	1.37	(0.25)	−0.03	(0.03)	1.03	(0.48)	3.91	(0.39)	0.17	(0.39)

The commonality across commodities is expressed by their loadings on market-wide and sector components, see also (4.3). Table 4.3 shows for each commodity level, slope, and curvature factor the estimated constant α, the loading on the market-wide component β, and the loading on the sector component γ. The α-parameters make sure that the idiosyncratic components have mean zero. All α level estimates are between 4 and 5, due to our standardization procedure, and are highly significant. The α estimates corresponding to slope and curvature are almost all not significantly different from zero. Two noteworthy exceptions are gold and cocoa. The negative α parameter for gold is in line with our expectations because its futures curve is often in contango.

All loadings on the market-wide level component (β's) are positive (or not significantly different from zero), which indicates that there exists a link between the levels of different commodity prices. The loadings on the market-wide slope component are positive for energy commodities, negative for most of the metal commodities, and close to zero for all other commodities. The loadings on the market-wide curvature component are all very different, ranging from -1.93 to 2.78. Most of them are not significantly different from zero. The loadings on the sector components give more insight in intra-sector commonality. In general, all sector loadings have the same sign within the corresponding sector, which implies commonality. The few slope and curvature loadings that have opposite signs are not significant. In line with previous results, all loadings on common components point in the direction of commonality.

The last column of Table 4.3 shows the variance estimates of the measurement equation errors. We assume that these variances are commodity specific but within each commodity they are the same for all different contract maturities. We believe this assumption is appropriate because the factor structure can already account for volatility differences along the term-structure dimension due to the time-to-maturity dependent factor loadings. Almost all estimated variances are well below the variances of the factor disturbances.

4.5.2 Importance of common factors

The loading estimates indicate that there is commonality across commodities. Using variance decompositions we investigate the importance of the common components. We decompose the variation in commodity level, slope, and curvature factors into parts driven by the market-wide, sector, and idiosyncratic components. As mentioned in Kose, Otrok, and Whiteman (2003) and Diebold, Li, and Yue (2008), the market-wide, sector, and commodity-specific components may be correlated as they are extracted from a finite sample. Hence we orthogonalize the extracted components using a Cholesky decomposition to ensure that they add up.[17] Then, we can use (4.3) to write

$$\begin{aligned}
\text{var}\left(\Delta l_{i,t}\right) &= \left(\beta_i^L\right)^2 \text{var}\left(\Delta L_{market-wide,t}\right) + \left(\gamma_i^L\right)^2 \text{var}\left(\Delta L_{sector,t}\right) + \text{var}\left(\Delta L_{i,t}\right), \\
\text{var}\left(s_{i,t}\right) &= \left(\beta_i^S\right)^2 \text{var}\left(S_{market-wide,t}\right) + \left(\gamma_i^S\right)^2 \text{var}\left(S_{sector,t}\right) + \text{var}\left(S_{i,t}\right), \\
\text{var}\left(c_{i,t}\right) &= \left(\beta_i^C\right)^2 \text{var}\left(C_{market-wide,t}\right) + \left(\gamma_i^C\right)^2 \text{var}\left(C_{sector,t}\right) + \text{var}\left(C_{i,t}\right).
\end{aligned} \tag{4.5}$$

For the level factors we decompose the variances of the first differenced series as the variance of a non-stationary series is undefined. The fraction of, e.g., the WTI crude oil level factor variance driven by the market-wide component is given by

$$\frac{\left(\beta_{WTI}^L\right)^2 \text{var}\left(\Delta L_{market-wide,t}\right)}{\text{var}\left(\Delta l_{WTI,t}\right)}.$$

The fractions of explained variance per component are shown in Table 4.4. The market-wide level component explains on average 23.3% of the variance of the commodity level factors, while the sector component explains 38.6%. However, the differences across commodities are large. For example, in the case of feeder cattle the market-wide component explains close to nothing of its level variation while for corn the market-wide component explains 72.1% of its variation. Overall, the market-wide component explains quite some variation of the silver, softs and grains levels. The energy sector component explains around 65% of all energy commodities, except natural gas. Similar observations can be made for the metal commodities (precious metals versus industrial metals) and the grain commodities (wheat versus corn and soybeans).

[17] We put the market-wide component first, followed by the sector component, and last the idiosyncratic component.

Table 4.4 Variance decompositions

The table presents the percentage of explained variation of the level, slope and curvature factors by the market-wide, sector, and idiosyncratic components. As the level factors are non-stationary, their analysis is done on first differences.

	ΔLevel			Slope			Curvature		
Commodity	Market	Sector	Idio.	Market	Sector	Idio.	Market	Sector	Idio.
Brent crude oil	13.5%	65.8%	20.7%	83.2%	7.5%	9.3%	0.1%	93.7%	6.2%
WTI crude oil	12.6%	66.6%	20.8%	85.1%	4.9%	10.0%	1.1%	90.4%	8.5%
Gasoil	12.3%	67.3%	20.5%	88.8%	0.1%	11.1%	28.8%	54.6%	16.6%
Heating oil	11.9%	66.9%	21.3%	69.8%	18.4%	11.7%	58.0%	9.5%	32.5%
Natural gas	11.2%	35.1%	53.7%	43.0%	38.4%	18.6%	34.5%	15.0%	50.5%
Gasoline	12.9%	66.3%	20.8%	77.8%	4.5%	17.7%	47.1%	0.8%	52.1%
Gold	17.9%	11.6%	70.5%	56.8%	37.4%	5.8%	-	-	-
Silver	40.6%	26.7%	32.7%	50.1%	32.7%	17.2%	-	-	-
Aluminum	10.2%	47.9%	42.0%	0.8%	90.7%	8.5%	-	-	-
Copper	12.3%	67.1%	20.6%	1.5%	92.2%	6.3%	-	-	-
Lead	10.6%	59.5%	29.8%	6.7%	84.4%	8.9%	-	-	-
Nickel	11.4%	64.9%	23.7%	6.3%	92.5%	1.2%	-	-	-
Zinc	10.0%	66.3%	23.7%	0.1%	99.4%	0.5%	-	-	-
Cocoa	34.6%	8.0%	57.4%	2.7%	2.8%	94.4%	2.6%	7.5%	89.9%
Coffee	32.7%	22.0%	45.4%	1.6%	90.7%	7.7%	16.8%	34.2%	49.1%
Cotton	44.3%	12.6%	43.1%	1.0%	9.0%	90.0%	2.3%	92.7%	5.0%
Sugar	24.6%	8.7%	66.7%	1.5%	3.8%	94.7%	69.2%	1.6%	29.3%
Corn	72.1%	1.2%	26.7%	0.0%	71.6%	28.4%	3.6%	85.4%	11.0%
Soybeans	68.2%	0.6%	31.2%	0.1%	44.6%	55.3%	34.1%	56.8%	9.1%
Chicago wheat	37.2%	41.9%	20.9%	0.0%	93.2%	6.8%	29.7%	63.4%	6.9%
Kansas wheat	36.3%	42.7%	21.0%	0.1%	87.0%	12.9%	34.2%	61.4%	4.4%
Feeder cattle	0.0%	42.4%	57.6%	0.0%	61.2%	38.8%	32.9%	28.8%	38.3%
Lean hogs	14.3%	8.7%	76.9%	21.5%	33.4%	45.1%	7.3%	68.6%	24.1%
Live cattle	7.7%	24.6%	67.7%	0.1%	65.3%	34.6%	6.1%	88.1%	5.8%

The variance decomposition results for the slope factor are again diverse. The market-wide, sector, and idiosyncratic components explain on average 24.9%, 48.6%, and 26.5%, respectively, of the commodities' slope factors. The market-wide component explains more than 70% of the energy commodities slope variation, with the exception of natural gas. Also the precious metals are driven by a common market component for more than 50%. All other slope factors are mostly explained by a sector component, except for cocoa, cotton, sugar and soybeans.

The market-wide curvature component explains on average 24.0% of the commodities' curvature factors, although there is no clear pattern in these results. The energy sector component is mostly tilted towards crude oils and gasoil. For the soft commodities the idiosyncratic component explain on average most of the variation, while for the grains and meats it is the sector curvature component.

The variance decomposition results indicate that 61.9% of the level variation, 73.5% of the slope variation and 74.2% of curvature variation is explained by common factors. In all three cases the market-wide component explains a quarter of the variation and the remainder is due to sector commonality.

4.5.3 Factor dynamics

Now we have established that there are common factors that drive commodities futures curves and that these common factors explain a large part of individual commodities variations, we focus on the dynamics of these factors. Table 4.5 shows the estimates related to the state equations in (4.4). Recall that we have assumed that both the autoregressive coefficient matrices and the covariance matrices are diagonal. Furthermore, the market-wide, sector and idiosyncratic level components are assumed to be non-stationary. We model their first differences as an AR(1) process. The first three columns in Table 4.5 show the autoregressive coefficients and the last three columns show the variances of the factor disturbances. Focusing on the market-wide and sector components, all ϕ estimates are positive (except for one insignificant grains sector parameter). The slope factors are more persistent than the curvature components. The parameter estimates for the idiosyncratic factors are in line with the corresponding common factors. Most of the level ϕ coefficients are not significantly different from zero, which implies that levels evolve as pure random walk processes. The AR parameters of the slope factors range between 0.61 and 0.99. The only exception is heating oil with an AR(1) parameter of -0.33. The curvature factors show similar coefficients as those of the slope factors but are slightly less persistent.

Table 4.5 Joint model factor dynamics

The table presents the dynamics of the unobserved states by showing the parameter estimates of the state equations. The autoregressive parameters are the diagonal elements of the Φ-matrices. All level factors are modeled in first differences due to their non-stationary behavior. The disturbance variances correspond to elements in Σ_{η_y}, which are multiplied by 1,000 for readability reasons. Note that the disturbance variances of the market-wide and sector components are fixed for identifications purposes. Note that we do not include a curvature factor for the metal commodities, which is represented by a horizontal dash.

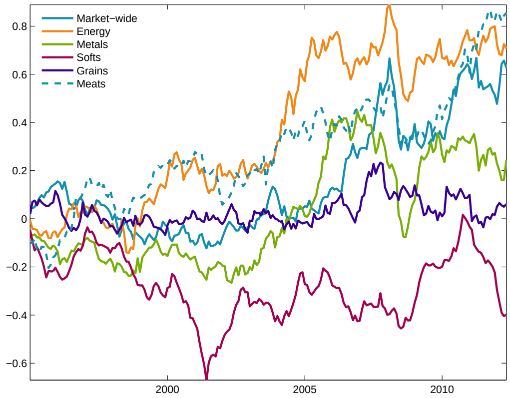
		Autoregressive parameters (ϕ)						Disturbance variances					
Sector	Factor	ΔLevel		Slope		Curv.		ΔLevel		Slope		Curv.	
	Market-wide	0.11	(0.05)	0.93	(0.46)	0.78	(0.26)	1.00	-	1.00	-	1.00	-
	Energy	0.12	(0.04)	0.79	(0.22)	0.88	(0.29)	1.00	-	1.00	-	1.00	-
	Metals	0.12	(0.05)	0.96	(0.94)	-	-	1.00	-	1.00	-	-	-
	Softs	0.45	(0.28)	0.93	(0.53)	0.87	(0.63)	1.00	-	1.00	-	1.00	-
	Grains	−0.12	(0.06)	0.94	(0.54)	0.88	(0.35)	1.00	-	1.00	-	1.00	-
	Meats	0.05	(0.05)	0.88	(0.38)	0.72	(0.18)	1.00	-	1.00	-	1.00	-
Energy	Brent crude oil	−0.07	(0.09)	0.92	(0.75)	0.83	(3.20)	0.11	(1.95)	0.20	(6.64)	0.03	(66.53)
	WTI crude oil	0.11	(0.07)	0.82	(0.32)	0.83	(0.23)	0.16	(1.39)	0.72	(0.97)	4.81	(0.23)
	Gasoil	−0.60	(0.15)	0.77	(0.20)	0.64	(0.14)	0.16	(1.46)	1.62	(0.50)	13.23	(0.19)
	Heating oil	−0.53	(0.16)	−0.33	(0.37)	0.83	(0.24)	0.08	(2.72)	0.28	(6.88)	12.52	(0.21)
	Natural gas	−0.03	(0.05)	0.91	(0.40)	0.88	(0.32)	3.92	(0.22)	19.60	(0.12)	53.52	(0.11)
	Gasoline	0.06	(0.06)	0.61	(0.12)	0.60	(0.11)	0.35	(0.88)	5.82	(0.21)	34.36	(0.12)
Metals	Gold	−0.09	(0.05)	0.98	(2.10)	-	-	1.73	(0.33)	299.85	(86.52)	-	-
	Silver	−0.24	(0.06)	0.99	(2.91)	-	-	4.96	(0.20)	133.71	(0.09)	-	-
	Aluminum	0.01	(0.06)	0.85	(0.40)	-	-	0.98	(0.52)	11.67	(0.21)	-	-
	Copper	0.09	(0.08)	0.97	(1.05)	-	-	1.10	(0.72)	34.54	(0.12)	-	-
	Lead	0.04	(0.05)	0.82	(0.23)	-	-	3.31	(0.26)	51.64	(0.09)	-	-
	Nickel	−0.02	(0.05)	0.86	(0.32)	-	-	5.23	(0.20)	205.89	(0.20)	-	-
	Zinc	−0.07	(0.07)	0.91	(0.53)	-	-	1.67	(0.45)	200.08	(0.25)	-	-
Softs	Cocoa	−0.14	(0.06)	0.96	(0.81)	0.80	(0.21)	5.05	(0.21)	2.80	(0.28)	8.37	(0.21)
	Coffee	−0.14	(0.07)	0.96	(0.76)	0.82	(0.22)	5.70	(0.23)	0.03	(31.22)	20.69	(0.15)
	Cotton	−0.12	(0.09)	0.96	(0.81)	0.82	(0.72)	2.17	(0.48)	7.85	(0.18)	6.87	(9.71)
	Sugar	0.09	(0.05)	0.88	(0.30)	0.85	(0.26)	3.77	(0.24)	13.97	(0.13)	28.52	(0.14)
Grains	Corn	0.04	(0.09)	0.92	(0.47)	0.76	(0.19)	0.66	(1.26)	5.82	(0.22)	16.20	(0.21)
	Soybeans	−0.19	(0.08)	0.92	(0.47)	0.81	(0.32)	1.14	(0.60)	4.40	(0.23)	7.71	(0.79)
	Chicago wheat	−0.25	(0.68)	0.89	(4.13)	0.88	(0.53)	0.01	(88.34)	0.03	(88.67)	2.79	(1.54)
	Kansas wheat	0.20	(0.09)	0.87	(0.32)	0.92	(1.63)	0.15	(2.77)	1.60	(0.61)	0.73	(12.67)
Meats	Feeder cattle	0.28	(0.21)	0.81	(0.43)	0.75	(0.18)	0.04	(24.97)	0.53	(5.93)	6.37	(0.28)
	Lean hogs	0.12	(0.05)	0.88	(0.32)	0.80	(0.22)	1.26	(0.47)	16.42	(0.15)	104.90	(0.13)
	Live cattle	0.29	(0.16)	0.82	(0.28)	0.69	(1.63)	0.02	(29.99)	1.94	(1.08)	0.18	(89.28)

The right-hand side of Table 4.5 shows the estimated variances of the factor disturbances. The variances of the common factor disturbances are fixed for identification purposes. The estimates of the idiosyncratic disturbance variance cannot be compared across commodities or factors because the magnitude of the factor loadings is not the same. This can be seen in Appendix 4.A Equation (4.A.1) where all idiosyncratic states are premultiplied by the matrix A, which contains the commodity specific factor loadings. For metal commodities the λ_i parameter is small, which results in close to zero loadings on the slope factor for most maturities and hence seemingly large error variances.

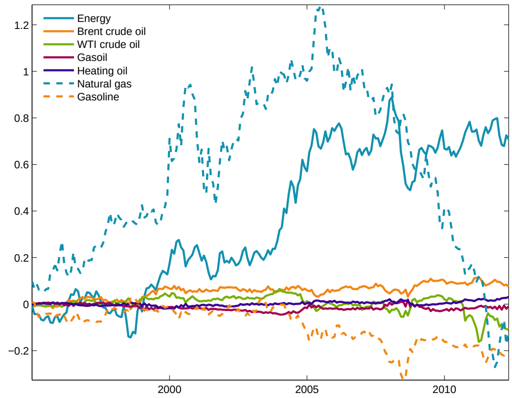
Figure 4.8 Joint model - extracted commodity level factors

These figures show the extracted level factors of our joint model estimated using all 24 commodities. Subfigure A shows the market-wide and sector components. Subfigures B-F show the estimated idiosyncratic level factors for the commodities of a specific sector.

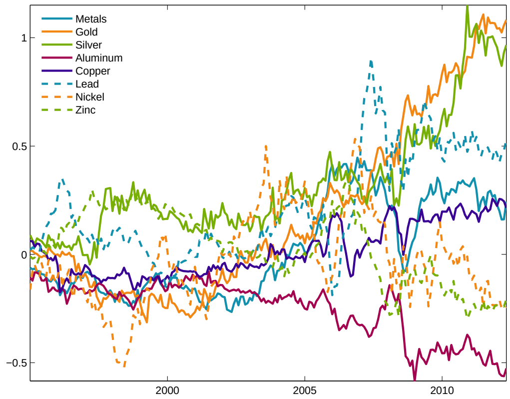
(a) Level - market-wide



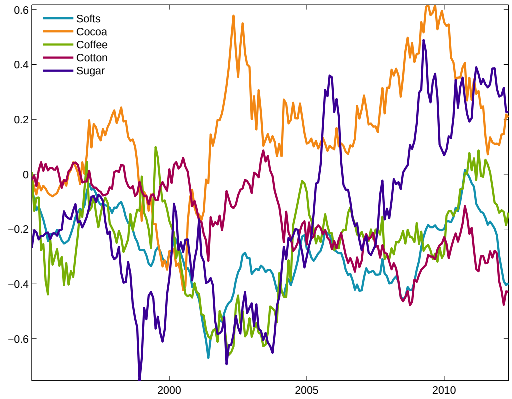
(b) Level - energy



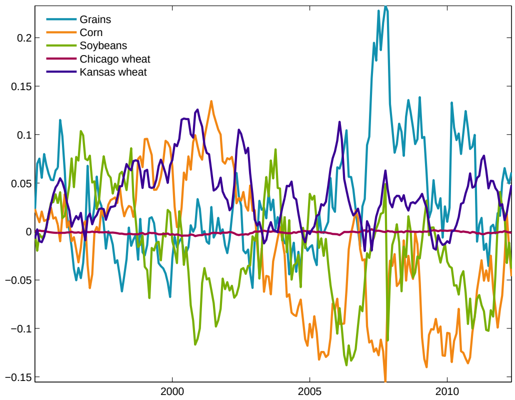
(c) Level - metals



(d) Level - softs



(e) Level - grains



(f) Level - meats

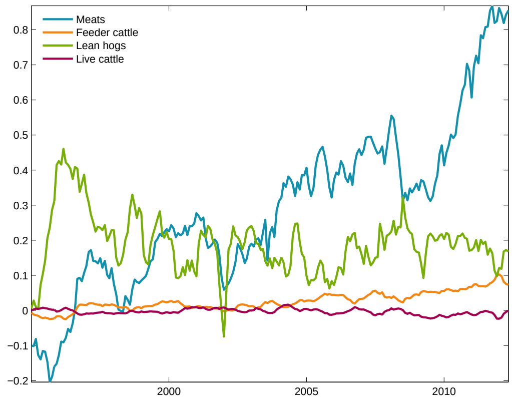
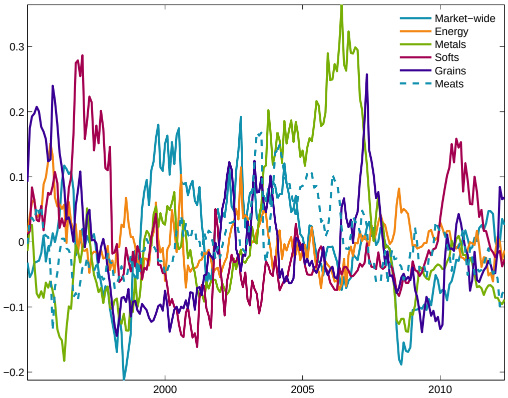


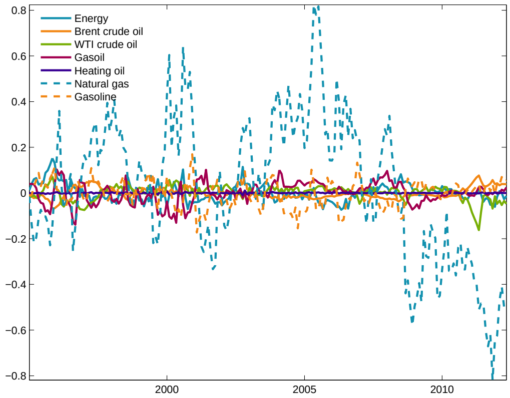
Figure 4.9 Joint model - extracted commodity slope factors

These figures show the extracted slope factors of our joint model estimated using all 24 commodities. Subfigure A shows the market-wide and sector components. Subfigures B-F show the estimated idiosyncratic slope factors for the commodities of a specific sector.

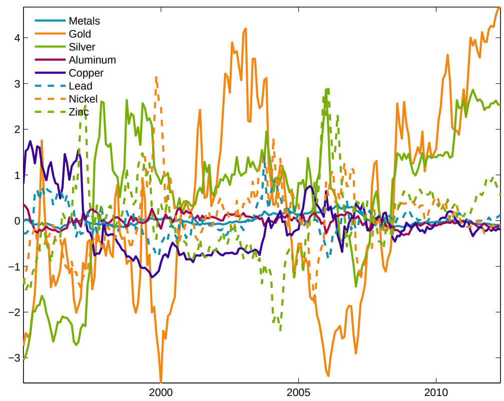
(a) Slope - market-wide



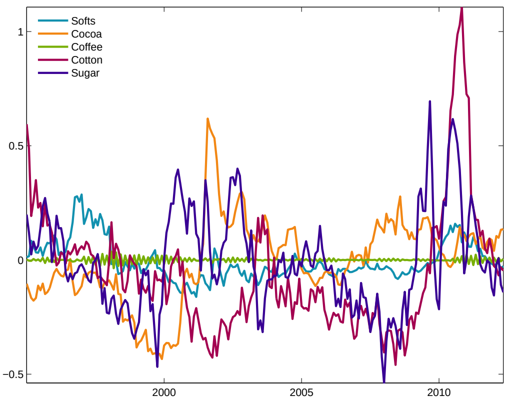
(b) Slope - energy



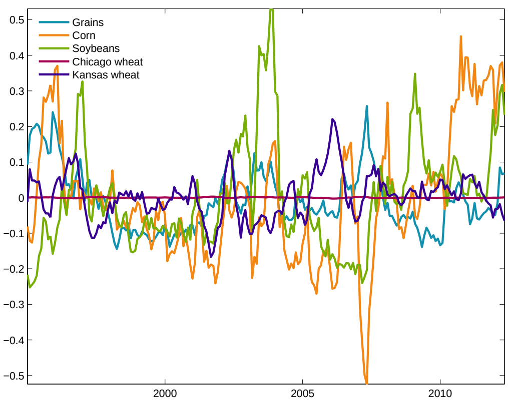
(c) Slope - metals



(d) Slope - softs



(e) Slope - grains



(f) Slope - meats

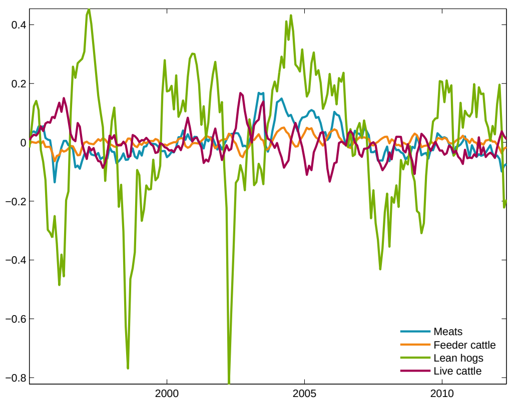
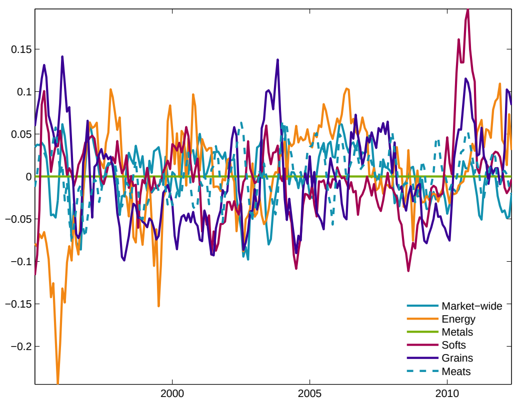


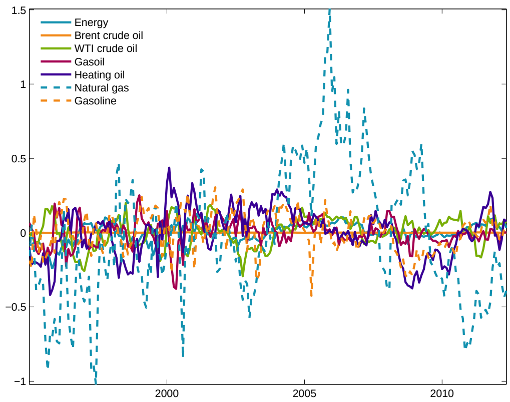
Figure 4.10 Joint model - extracted commodity curvature factors

These figures show the extracted curvature factors of our joint model estimated using all 24 commodities. Subfigure A shows the market-wide and sector components. Subfigures B-E show the estimated idiosyncratic curvature factors for the commodities of a specific sector. Note that the metal commodities subfigure is missing because we use a two factor model for the metal commodities.

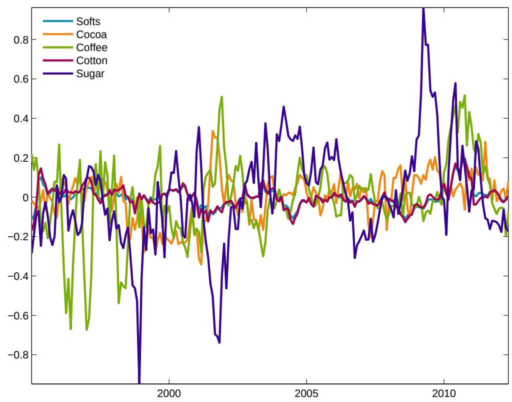
(a) Curvature - market-wide



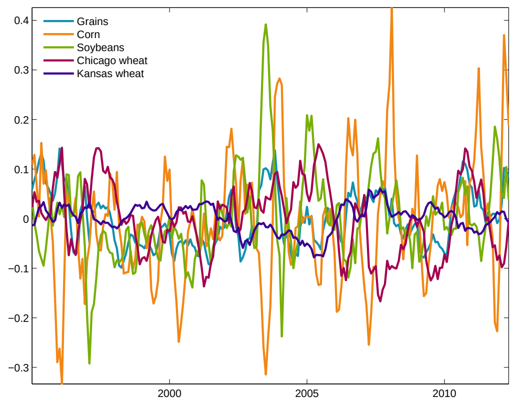
(b) Curvature - energy



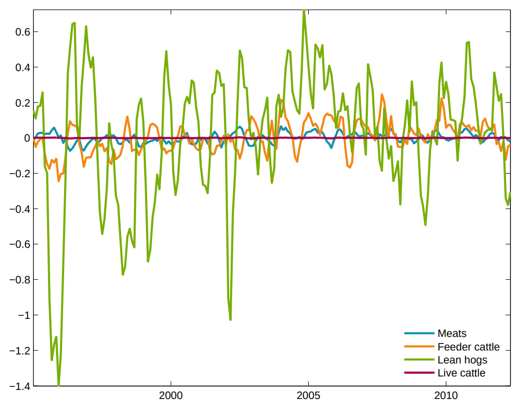
(c) Curvature - softs



(d) Curvature - grains



(e) Curvature - meats



Figures 4.8, 4.9, and 4.10 show the unobserved level, slope, and curvature factors and their different components. In Subfigure A of each figure we show the market-wide component together with the five sector components. In Subfigures B-F we show the sector components together with the corresponding idiosyncratic commodity components. The market-wide level factor shows an increase in 2006-2007 and 2011, in line with patterns observed in all major commodity indices. The energy and metal sectors show similar behavior although the factors peak at different points in time, in line with subindices like GSCI Energy, Precious Metals and Industrial Metals. The other subfigures show that the scale of the idiosyncratic level factors varies across commodities but is in line with the size of the estimated disturbance variances. The slope factors in Figure 4.9 show quite some variability. They are both negative and positive, which indicates that the futures curves interchange between contango and backwardation. The scale of the metals' slope factors is large compared to the other commodities due to the λ-dependent loading correction we have introduced in (4.2). Last, the curvature factors in Figure 4.10 show more mean-reversion compared to the slope factors.

4.5.4 Economic interpretation of unobserved states

We have established that there are common factors in commodities futures curves and these factors explain a substantial part of the variation in level, slope and curvature factors that drive observed futures prices. As these common components are unobserved, it is not straightforward which economic mechanism is underlying this. In this section we link our unobserved common factors to observed macroeconomic variables.

Existing literature provides us with a range of variables that could be related to our level, slope or curvature factors. The theory of normal backwardation (Keynes, 1930) argues that commodity producers and inventory holders hedge their risk by shorting futures. To induce risk-averse speculators into taking the opposite long positions, current futures prices are set at a discount (i.e., are "backwardated") to expected future spot prices at maturity. Therefore the ratio of hedgers versus speculators could be related to the shape of the futures curve and hence to our factors. Alternatively, the theory of storage (Kaldor, 1939; Working, 1949) argues that convenience yield, basis, and inventories are closely related. In our set-up we capture the convenience yield in our slope (see Section 4.4.2) factors. Besides commodity specific variables, we also use macroeconomic variables to interpret our factors.

We collect a large database of macroeconomic and commodity specific variables. We consider the same set of 108 macroeconomic variables as in Stock and Watson (2012). Following Stock and Watson (2012) we transform the variables to ensure stationarity and assign them to 12 categories: GDP components, industrial production, employment, unemployment rate, housing, business inventories, wages, interest rates, money, exchange rates, stock prices, and consumer expectations.[18] For each category we apply a Principal Components Analysis (PCA) to summarize the variables within that group and proceed with the principal component that explains most of their variation. In this way we reduce the dimension to 12 series that all correspond to a particular macroeconomic category. Besides the Stock and Watson (2012) dataset, we add a "financial conditions" group to capture investor expectations and market conditions. This category consists of the Aruoba, Diebold, and Scotti (2009) (ADS) business conditions index, the Baker and Wurgler (2006) sentiment index (Lutzenberger, 2014), and the Baltic dry shipping index (Bakshi et al., 2011).[19] Even though our financial conditions variable is based on macroeconomic series that are already included in other categories, it is not highly correlated with these other explanatory variables. We collect commodity inventory and hedging pressure data following the methodology of Gorton, Hayashi, and Rouwenhorst (2013). Last, we add volatility of the Commodity Research Bureau (CRB) spot market price index (Pindyck, 2004) as possible candidate. The in total 209 collected individual series result in 17 stationary variables. A complete overview of the individual series, their sources, categories and transformations is given in Appendix 4.E.

To determine which variables are most related to which common component we use a multivariate regression. Table 4.6 shows the variables that have a statistically significant coefficient. Up to one third of the variation of the differenced level components can be explained by our explanatory variables. These percentages are higher for our slope and curvature components.

Focusing on Panel A, the results for the common market-wide components are in line with our expectations. Changes in the level component, i.e. returns, are related to returns in

[18] We leave out the Prices category due to endogeneity issues.

[19] The ADS index is designed to track real business conditions at high frequency. Baker and Wurgler (2006) define sentiment as investor propensity to speculate. The Baltic Dry Index is an indicator of transportation costs for raw materials shipped by sea.

foreign exchange and equity markets. As all commodity futures contracts we examine are denominated in dollars, a stronger dollar leads to lower commodity prices, which is reflected by the negative coefficient. The positive relation between equity and commodity prices is surprising as commodities are often used for diversification purposes (see e.g. Erb and Harvey, 2006; Gorton and Rouwenhorst, 2006). However it is in line with more recent findings of Singleton (2014) and Tang and Xiong (2012). The (almost significant) positive relation with hedging pressure is in line with the theory of normal backwardation Keynes (1930) and the results in Gorton, Hayashi, and Rouwenhorst (2013); Hamilton and Wu (2014).[20]

[20]Note that Gorton, Hayashi, and Rouwenhorst (2013) Table XI reports significant negative slope coefficients. However, they define hedging pressure as net long positions of commercials (hedgers) while we look at, the opposite, net short positions of hedgers.

Table 4.6 Interpretation common components

The table presents observed macroeconomic and commodity specific variables that are related to our unobserved states. Each component is regressed on a set of 17 explanatory variables. The table presents the statistically significant coefficients with their corresponding t-stat. The macroeconomic variables are taken from Stock and Watson (2012), while the commodity specific variables are collected as described by Gorton et al. (2013). Full details on the data series are given in Appendix 4.E.

ΔLevel		Slope		Curvature	
Variable	t-stat	Variable	t-stat	Variable	t-stat
Panel A Market-wide components					
$R^2 = 16.5\%$		$R^2 = 60.1\%$		$R^2 = 22.6\%$	
Equity	3.29	Business inventories	4.43	Business inventories	2.53
Exchange rates	−2.48	Employment	−2.74	Hedging pressure	−3.39
Hedging pressure	1.93	Financial conditions	8.29	Interest rates	3.76
		Hedging pressure	10.04		
		Housing	8.54		
		Industrial production	−2.78		
Panel B Energy sector components					
$R^2 = 31.0\%$		$R^2 = 24.3\%$		$R^2 = 31.1\%$	
Employment	−2.49	Business inventories	−4.63	Business inventories	3.74
Equity	2.09	Commodity inventories	−3.02	Financial conditions	4.90
Exchange rates	−2.05	Financial conditions	−5.10	Hedging pressure	−4.75
Financial conditions	2.80			Industrial production	−3.55
Hedging pressure	6.92			Interest rates	1.99
Housing	3.69			Unemployment	−2.23
Interest rates	2.34				
Panel C Metals sector components					
$R^2 = 29.9\%$		$R^2 = 49.1\%$			
Employment	−3.64	Business inventories	7.93		
Equity	4.72	Commodity inventories	−2.09		
Exchange rates	−3.36	Hedging pressure	−7.53		
Industrial production	2.40	Housing	8.41		
Interest rates	2.43				
Wages	−2.68				
Panel D Softs sector components					
$R^2 = 36.9\%$		$R^2 = 47.0\%$		$R^2 = 41.7\%$	
Employment	−4.00	Employment	6.84	Business inventories	2.53
Equity	2.00	Hedging pressure	5.10	Commodity volatility	−2.06
Hedging pressure	8.02	Housing	−4.83	Employment	4.99
Housing	3.09			Housing	−5.12
				Industrial production	−2.28
Panel E Grains sector components					
$R^2 = 7.1\%$		$R^2 = 56.1\%$		$R^2 = 60.8\%$	
		Business inventories	−2.90	Employment	4.36
		Commodity volatility	−2.67	Financial conditions	−5.20
		Employment	2.23	Hedging pressure	13.26
		Financial conditions	−7.82	Industrial production	−2.32
		Hedging pressure	10.62		
		Housing	3.74		
		Interest rates	−3.20		
Panel F Meats sector components					
$R^2 = 19.0\%$		$R^2 = 39.8\%$		$R^2 = 10.9\%$	
Hedging pressure	4.82	Business inventories	4.66	Hedging pressure	2.14
		Employment	−4.44		
		Hedging pressure	3.27		
		Housing	8.43		
		Wages	−2.88		

Business inventories (new orders growth), housing (construction growth) and financial conditions are all considered leading economic indicators. Their positive relation with our market-wide slope factor implies a shift to more backwardated commodities curves when economic future perspectives are positive.[21] The same holds for hedging pressure. When it goes up (more net short positions by hedgers), our slope component goes up, which results in a backwardated futures curve. The relations between the market-wide slope and industrial production growth and changes in employment are negative, hence the commodity curves are more backwardated at times when the economy slows down. This is not in line with our expectations, e.g. Fama and French (1988) find that when metal inventories are high, convenience yields are lower and the curve will be in contango.

The curvature component is related to business inventories, hedging pressure and interest rates. An increase in the curvature component leads to an increase in the price of mid-term contracts, while the contracts with very short or very long time to maturities are less affected. This seems to be coinciding with higher Treasury yields and lower hedging pressures.

Panels B to F show for each sector to which variables the components are linked. For the commodity specific variables we use only data of the commodities included in the sector of interest. In general, all level components are related to exchange rates, equity and hedging pressure. The grains level component is an exception as it hardly explained by any of our explanatory variables, shown by the R^2 of 7.1%. The sector slope components relate positively to hedging pressure, just as the market-wide slope component. Only for the metals sector the coefficient for hedging pressure is negative, which is opposite of what we would expect. It is also interesting to see that two sector slope components have a negative coefficient for commodity inventories. This is in line with the theory of storage as shown by, e.g., Gorton et al. (2013); Geman and Nhuyen (2005). Last, the sector curvature components show similar results as the market-wide curvature component. Business inventories, hedging pressure and yield curve variables have significant explanatory power. Furthermore, industrial production is negatively related to most sector curvature components, while employment variables are positively related.

Concluding, our unobserved common components relate to observed macroeconomic and commodity specific variables. In general, changes in the level components are related to

[21] Recall that the loadings on our slope factors are convex, hence a positive (negative) slope factor implies a backwardated (contangoed) futures curve.

equity, exchange rates and hedging pressure which is in line with results of existing literature. Also the slope components show significant coefficients for hedging pressure, housing, and commodity inventories that are in line with our expectations. Last, our curvature factor, which to the best of our knowledge has not yet been investigated in existing research, is positively related to interest rates and business inventories and negatively to industrial production.

4.5.5 Time variation

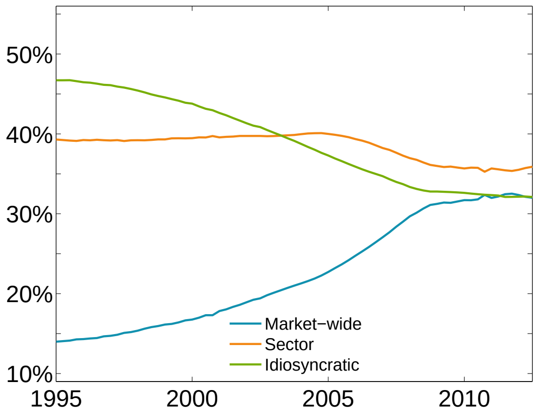
According to the ongoing debate on the financialization of commodities markets, the entrance of financial investors (around 2004-2005) has changed the commodities market dynamics (Cheng and Xiong, 2014). All our analyses so far are based on the full sample period from 1995 to 2012. It would be interesting to see if the amount of commonality varies over time during this period.

To investigate variation over time, we re-estimate the model attaching weights to the likelihood contributions of different observations in such a way that we emphasize specific data periods. The likelihood contribution of the observation at time $\Theta + k$ is given weight $\delta^{|k|}$, for $k = \ldots, -2, -1, 0, 1, 2, \ldots$, with $0 < \delta < 1$, resulting in estimates centered at $t = \Theta$. Repeating this for $\Theta = 1, 2, \ldots, T$ yields a sequence of smoothly time-varying parameter estimates. Although by design, this weighting will not produce any abrupt change, it nevertheless provides information about the presence, or otherwise, of temporal variation. We favor this approach to the use of subsamples because our data covers a limited time period and we do not want to impose breakdates ourselves. We use $\delta = 0.99$ to ensure that each estimate reflects information in a sample of reasonable effective size.

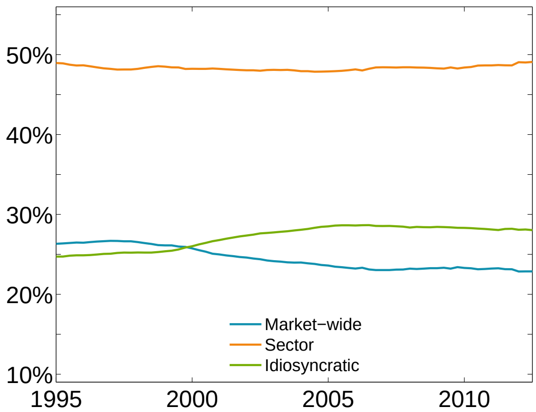
Figure 4.11 Moving window variance decompositions

These figures show the percentage of explained variation per component. Time-variation is introduced by estimating the model using a rolling window and applying a variance decomposition analysis at each point in time. The time-varying specification is estimated by attaching lower weights to more distant observations. The observation at time $\Theta + k$ is given weight $\delta^{|k|}$, for $k = \ldots, -2, -1, 0, 1, 2, \ldots$, with $0 < \delta < 1$, resulting in estimates centered at $t = \Theta$.

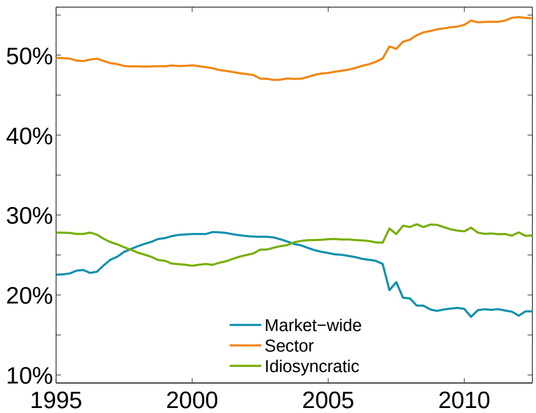
(a) ΔLevel



(b) Slope



(c) Curvature



Of particular interest is the relative importance over time of the common, market-wide and sector components. Figure 4.11 shows results of time-varying variance decompositions, which are at each point in time based on estimation results of the above described methodology. Both the lines corresponding to sector and idiosyncratic explained variation are based on averages across all sectors or commodities, respectively. Subfigure A shows that the market-wide component becomes increasingly important over time. In 1995 this component common to all commodities explained just 14% of total variance, while this percentage steadily increases to 32% in 2008 after which it stays constant. Most of the increase happened at the expense of the idiosyncratic components. Subfigure B shows the results for the slope factors. All three lines are pretty constant over time. Subfigure C shows slightly more variation over time for the curvature components. From 2007 onward, the sector components explain a larger part of total curvature variation, while the market-wide component explains less variation. Still, the total amount of curvature variation explained by common components remains quite constant over time.

Our findings show that only the level factors show an increase in commonality over time. These findings are in line with Tang and Xiong (2012).

4.5.6 Importance curve data

In all our analysis we use term-structure information while investigating commonality. An obvious advantage of this additional data is that it allows us to investigate commonality in curve shapes, besides the often investigated commonality in levels. We argue that the inclusion of more distant futures data not only gains additional insight but also affects the common level component analysis. To quantify potential differences due to the inclusion of term structure information, we redo some of our analysis with only first nearby contract information. This allows us to compare the market-wide level factors that we obtained using our full model and dataset with the market-wide level factor that we obtain from a restricted version of our model and a limited dataset.

Our one factor model based on only first nearby contracts is obtained by restricting (4.2)-(4.4) into

$$f_{i,t} = l_{i,t} + \nu_{i,t} = \alpha_i^L + \beta_i^L L_{market,t} + \gamma_i^L L_{sector,t} + L_{i,t} + \nu_{i,t}, \tag{4.6}$$

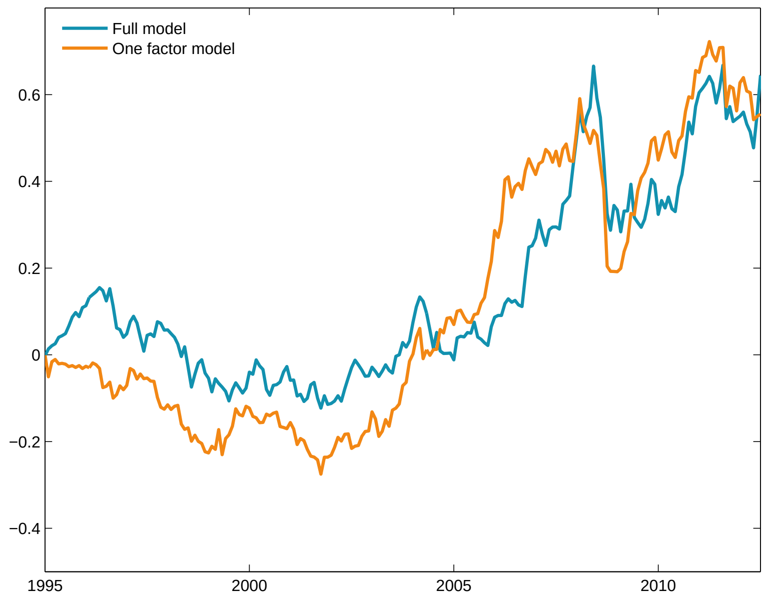
where

$$\Delta L_{x,t} = \phi^x \Delta L_{x,t-1} + \eta^L_{x,t}. \tag{4.7}$$

We estimate this model using the same Kalman filter methodology but use only data on first nearby contracts. Figure 4.12 shows a comparison of the extracted market-wide factors. The results from both methodologies are in line but there are some notable differences. First, the factor from the full model is pretty constant in the period 2004-2006 and 2009-2010, while the factor from the one-factor model rises sharply. However, after both subperiods, the full model factor quickly picks up and rises to the same level as the one-factor model. The differences in behavior are also reflected by the correlation of 0.53 between the first differenced series of both factors.

Figure 4.12 Comparison market-wide level components

Each line shows the market-wide level component of a particular model. The blue line corresponds to the full joint model with slope and curvature factors and curve information. The orange line is a restricted version of the full model with only level factors and first nearby contract information.



The extracted market-wide factors are just one side of the story. It is also interesting to investigate the other side, namely the loadings on these factors. Of most interest is the amount of commonality, i.e. the interplay of loadings and common factors. When we redo the variance decompositions we find that the market-wide factor explains a larger part of the variance in the one-factor model, 30.3% versus 23.3% for the full model. However, the seasonal commodities are less well explained compared to including the full dataset. The market-wide component explains 15.2% of their variation in the one-factor case, while this is 27.2% in our main results.

To see whether it is a coincidence that the seasonal commodities are better explained by our full curve model, we correct the first nearby contract data for seasonal effects. We do this by applying the same seasonal correction (based on our individual model estimates from Section 4.4.1). Note that in this way we favor the one-factor model as this correction is based on full curve data. Unreported results show that all previous reported findings hardly change. In other words, the seasonal correction has only a small effect on the difference between full curve and first contract results. This implies that there is a large difference between quantifying the commodity price levels based on only the first nearby contract data versus the levels based on the full commodities curve. Furthermore, commonality seems to be larger when only the front contracts are used. This was not expected as front contracts are more volatile than contracts further down the curve (see Table 4.1 and Samuelson 1965). A possible explanation is the effect of commodity indices. These indices (or funds that track them) are mostly invested in front contracts, which could lead to increased commonality (Tang and Xiong, 2012).

4.6 Conclusion

We investigate comovement across commodities by examining the commonality in the price levels and shapes of their futures curves. We use an enhanced version of the Nelson and Siegel (1987) model and extend the framework of Diebold, Li, and Yue (2008) to extract the factors that drive the individual commodity futures curves. Comovement across commodities is investigated by decomposing each individual factor in a market-wide, sector, and idiosyncratic component.

Using a monthly dataset of 24 commodities that are part of the S&P Goldman Sachs Commodity Index (GSCI) we show that there is comovement across commodity futures curves, either due to a market-wide or due to a sector component. Sector components explain close to 50% of the variation of our shape factors (slope and curvature), while a market-wide component explains 24% of their variations. For the individual commodity level factors, the percentage of explained variation by common (market-wide and sector) components is lower and on average 62%. Concluding, common components explain between 62% and 74% of the variation of individual commodities. In all cases the commonality is mostly driven by the sector components. For the shape related factors we find almost no variation in results over time. In contrast to the level factors where the market-wide component explains more variation over time. The percentage of explained variation starts at 14% in 1995 and increases to 32% in 2012.

The unobserved common components relate to macroeconomic and commodity specific variables in ways which are consistent with existing literature. Our level components relate to equity markets, exchange rates and hedging pressure. The slope components are linked to hedging pressure (theory of normal backwardation) and commodity inventories (theory of storage). Last, the newly introduced curvature components related to the yield curve, business inventories and industrial production.

The presented framework provides a way to include more futures data to investigate commonality across commodities. Using this framework we show that it is important to include the term-structure dimension in the analysis of comovement as it alters the findings on the extent of comovement. The current findings are insightful for portfolio construction, risk management and hedging purposes using commodity futures.

Appendix

4.A State space representation

The state space representation follows naturally from the model given by (4.2)-(4.4). The measurement equation in (4.A.1) is a combination of (4.2) and (4.3). Note that the individual latent level $l_{i,t}$, slope $s_{i,t}$, and curvature $c_{i,t}$ factors do not appear in the measurement equation, as we can link the observed futures prices $f_{i,t}(\tau)$ directly to the unobserved market-wide, sector and idiosyncratic components.

$$\begin{pmatrix} f_{1,t}(\tau_1) \\ f_{1,t}(\tau_2) \\ \vdots \\ f_{1,t}(\tau_{J_1}) \\ \vdots \\ f_{N,t}(\tau_{J_N}) \end{pmatrix} = A \begin{pmatrix} \alpha_1^L \\ \alpha_1^S \\ \alpha_1^C \\ \vdots \\ \alpha_N^C \end{pmatrix} + B \begin{pmatrix} L_{market,t} \\ S_{market,t} \\ C_{market,t} \end{pmatrix} + C \begin{pmatrix} L_{Energy,t} \\ S_{Energy,t} \\ C_{Energy,t} \\ \vdots \\ L_{Meats,t} \\ S_{Meats,t} \\ C_{Meats,t} \end{pmatrix} + A \begin{pmatrix} L_{1,t} \\ S_{1,t} \\ C_{1,t} \\ \vdots \\ C_{N,t} \end{pmatrix} + D \begin{pmatrix} \kappa_1 \\ \kappa_2 \\ \vdots \\ \kappa_N \end{pmatrix} + \begin{pmatrix} \nu_{1,t}(\tau_1) \\ \nu_{1,t}(\tau_2) \\ \vdots \\ \nu_{1,t}(\tau_{J_1}) \\ \vdots \\ \nu_{N,t}(\tau_{J_N}) \end{pmatrix} \tag{4.A.1}$$

where J_i is the number of available contracts of commodity i. As discussed in Section 4.2 J_i varies over time, yet for readability reasons we keep writing J_i instead of J_{i_t}.

$$A = \begin{pmatrix} 1 & \left(\frac{1-e^{-\lambda_1\tau_1}}{\lambda_1\tau_1}\right) & \left(\frac{1-e^{-\lambda_1\tau_1}}{\lambda_1\tau_1} - e^{-\lambda_1\tau_1}\right) & 0 & \cdots & \cdots & 0 \\ 1 & \left(\frac{1-e^{-\lambda_1\tau_2}}{\lambda_1\tau_2}\right) & \left(\frac{1-e^{-\lambda_1\tau_2}}{\lambda_1\tau_2} - e^{-\lambda_1\tau_2}\right) & 0 & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & \left(\frac{1-e^{-\lambda_N\tau_{J_N}}}{\lambda_N\tau_{J_N}}\right) & \left(\frac{1-e^{-\lambda_N\tau_{J_N}}}{\lambda_N\tau_{J_N}} - e^{-\lambda_N\tau_{J_N}}\right) \end{pmatrix},$$

$$B = \begin{pmatrix} \beta_1^l & \beta_1^s\left(\frac{1-e^{-\lambda_1\tau_1}}{\lambda_1\tau_1}\right) & \beta_1^c\left(\frac{1-e^{-\lambda_1\tau_1}}{\lambda_1\tau_1} - e^{-\lambda_1\tau_1}\right) \\ \beta_1^l & \beta_1^s\left(\frac{1-e^{-\lambda_1\tau_2}}{\lambda_1\tau_2}\right) & \beta_1^c\left(\frac{1-e^{-\lambda_1\tau_2}}{\lambda_1\tau_2} - e^{-\lambda_1\tau_2}\right) \\ \vdots & \vdots & \vdots \\ \beta_N^l & \beta_N^s\left(\frac{1-e^{-\lambda_N\tau_{J_N}}}{\lambda_N\tau_{J_N}}\right) & \beta_N^c\left(\frac{1-e^{-\lambda_N\tau_{J_N}}}{\lambda_N\tau_{J_N}} - e^{-\lambda_N\tau_{J_N}}\right) \end{pmatrix},$$

$$C = \begin{pmatrix} \gamma_1^l & \gamma_1^s\left(\frac{1-e^{-\lambda_1\tau_1}}{\lambda_1\tau_1}\right) & \gamma_1^c\left(\frac{1-e^{-\lambda_1\tau_1}}{\lambda_1\tau_1} - e^{-\lambda_1\tau_1}\right) & 0 & \cdots & \cdots & 0 \\ \gamma_1^l & \gamma_1^s\left(\frac{1-e^{-\lambda_1\tau_2}}{\lambda_1\tau_2}\right) & \gamma_1^c\left(\frac{1-e^{-\lambda_1\tau_2}}{\lambda_1\tau_2} - e^{-\lambda_1\tau_2}\right) & 0 & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \gamma_N^l & \gamma_N^s\left(\frac{1-e^{-\lambda_N\tau_{J_N}}}{\lambda_N\tau_{J_N}}\right) & \gamma_N^c\left(\frac{1-e^{-\lambda_N\tau_{J_N}}}{\lambda_N\tau_{J_N}} - e^{-\lambda_N\tau_{J_N}}\right) \end{pmatrix},$$

$$D = \begin{pmatrix} \cos\left(\omega g_1\left(t,\tau_1\right) - \omega\theta_1\right) & 0 & \cdots & 0 \\ \cos\left(\omega g_1\left(t,\tau_2\right) - \omega\theta_1\right) & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \cos\left(\omega g_1\left(t,\tau_{J_1}\right) - \omega\theta_1\right) & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cos\left(\omega g_N\left(t,\tau_{J_N}\right) - \omega\theta_N\right) \end{pmatrix}.$$

The transition equations of the latent states are given by (4.4). The market-wide, sector, and idiosyncratic components are assumed to have first-order autoregressive dynamics. For completeness, we also present them here:

$$\begin{pmatrix} \Delta L_{y,t} \\ S_{y,t} \\ C_{y,t} \end{pmatrix} = \begin{pmatrix} \phi_{11}^y & \phi_{12}^y & \phi_{13}^y \\ \phi_{21}^y & \phi_{22}^y & \phi_{23}^y \\ \phi_{31}^y & \phi_{32}^y & \phi_{33}^y \end{pmatrix} \begin{pmatrix} \Delta L_{y,t-1} \\ S_{y,t-1} \\ C_{y,t-1} \end{pmatrix} + \begin{pmatrix} \eta_{y,t}^L \\ \eta_{y,t}^S \\ \eta_{y,t}^C \end{pmatrix}, \tag{4.A.2}$$

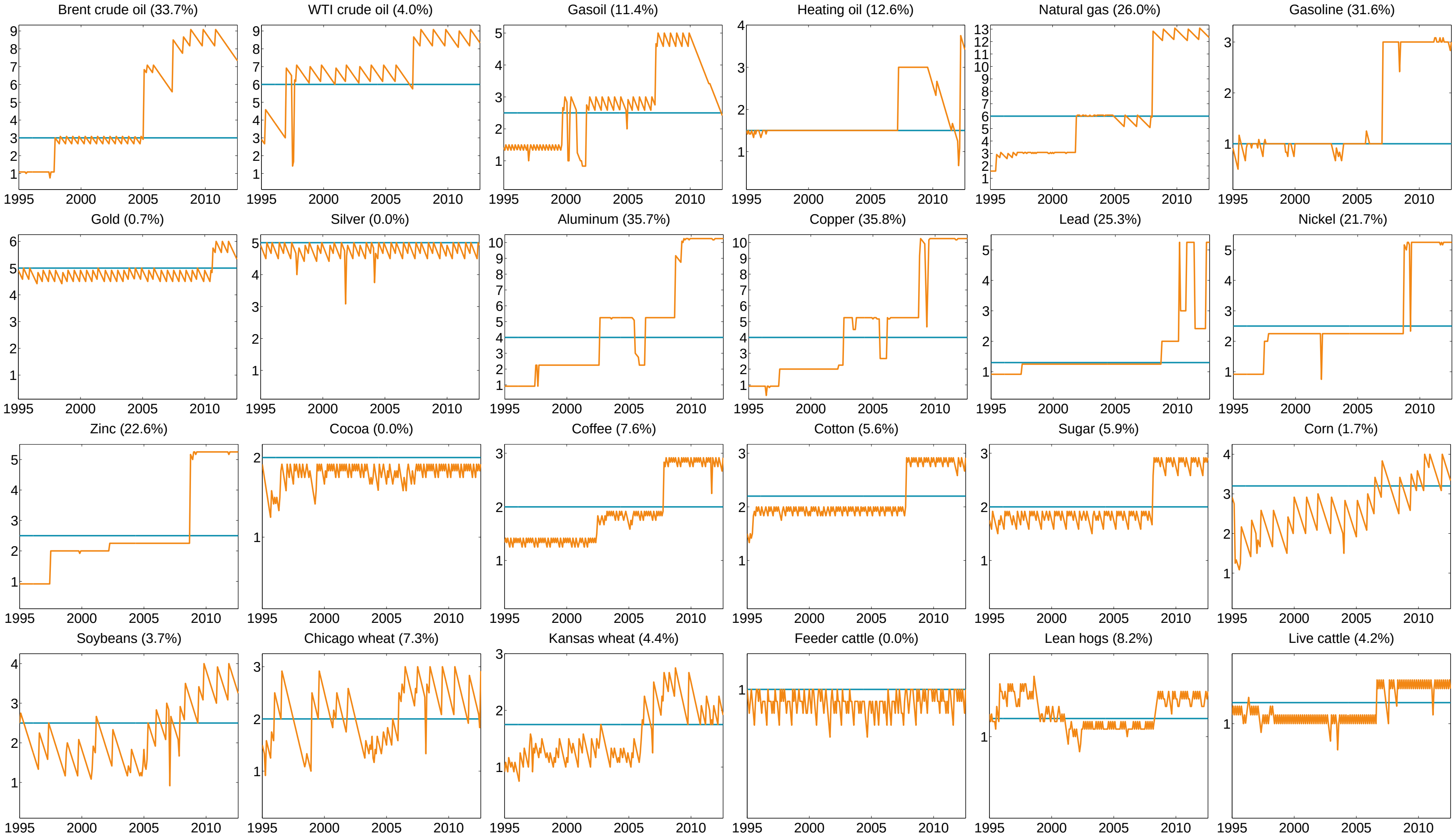
where $y = \{\textit{market, sector, idiosyncratic}\}$, and the disturbances $\eta_{y,t} = (\eta_{y,t}^L, \eta_{y,t}^S, \eta_{y,t}^C)$ are normally distributed with covariance matrix Σ_{η_y}.

4.B Maturity bound

We choose to introduce a maturity bound to exclude long-dated contracts. By limiting the term-structure dimension variation within a commodity, we can keep using a fixed λ value. Furthermore, these long-dated contracts are possibly less liquid and hence have more noisy price information, which could otherwise affect our results. The introduced maturity bound excludes on average 10% of our observations. Figure 4.B.1 gives more details on individual commodities. Note that the longest available maturity varies substantially over time.

Figure 4.B.1 Maturity bounds

These figures show for each commodity the maturity bound and the longest available maturity (in years).



4.C Additional individual commodity results

Detailed individual model results

This section provides the estimation results for all individual model specifications. Based on these results we decide on the final model specification for every commodity. The final model choice is based on several criteria. First, we compare the results of three factor models with and without seasonal term, to see if the exposure κ to the seasonal correction is significantly different from zero. There are 13 commodities in Table 4.C.1 that have highly significant κ parameters, ranging between 0.73 and 7.78. There is a big gap between the 11 commodities which have a t-stat below 5, and the remaining 13 commodities which have a t-stat above 400. Therefore we decide to not include a seasonal correction when κ is below 0.1. Second, for the non-seasonal commodities, we need to decide if we include a third curvature factor or not. Non-reported AIC or BIC values indicate that we should always choose for the larger models. However, the λ values for some metal commodities in Table 4.C.1 are below 0.5 which leads to slope and curvature loadings that are close to opposite with a correlation below -0.80. Therefore we decide to exclude the curvature factor for all metal commodities.

Principal component analysis on raw prices

To provide additional evidence that the Nelson and Siegel (1987) model is suited for commodity futures, we apply PCA to raw price data. In order to apply PCA, we need a balaced data panel. Therefore we exclude contract data if they have missing price data for more than 10% of the time periods. Then, for this selection, we exclude months where one of the contracts has missing price data. In the end we are left with a balanced sample with no missing observations.

Figure 4.C.2 shows the loadings of the first three principal components (based on the covariance matrix). For most commodities these loadings resemble the level, slope and curvature factor loadings. The only exceptions are some commodities with a pronounced seasonal pattern, namely heating oil, natural gas, gasoline, cotton and lean hogs. Due to the Samuelson (1965) effect, results based on the correlation matrix are even more similar to level, slope and curvature loadings.

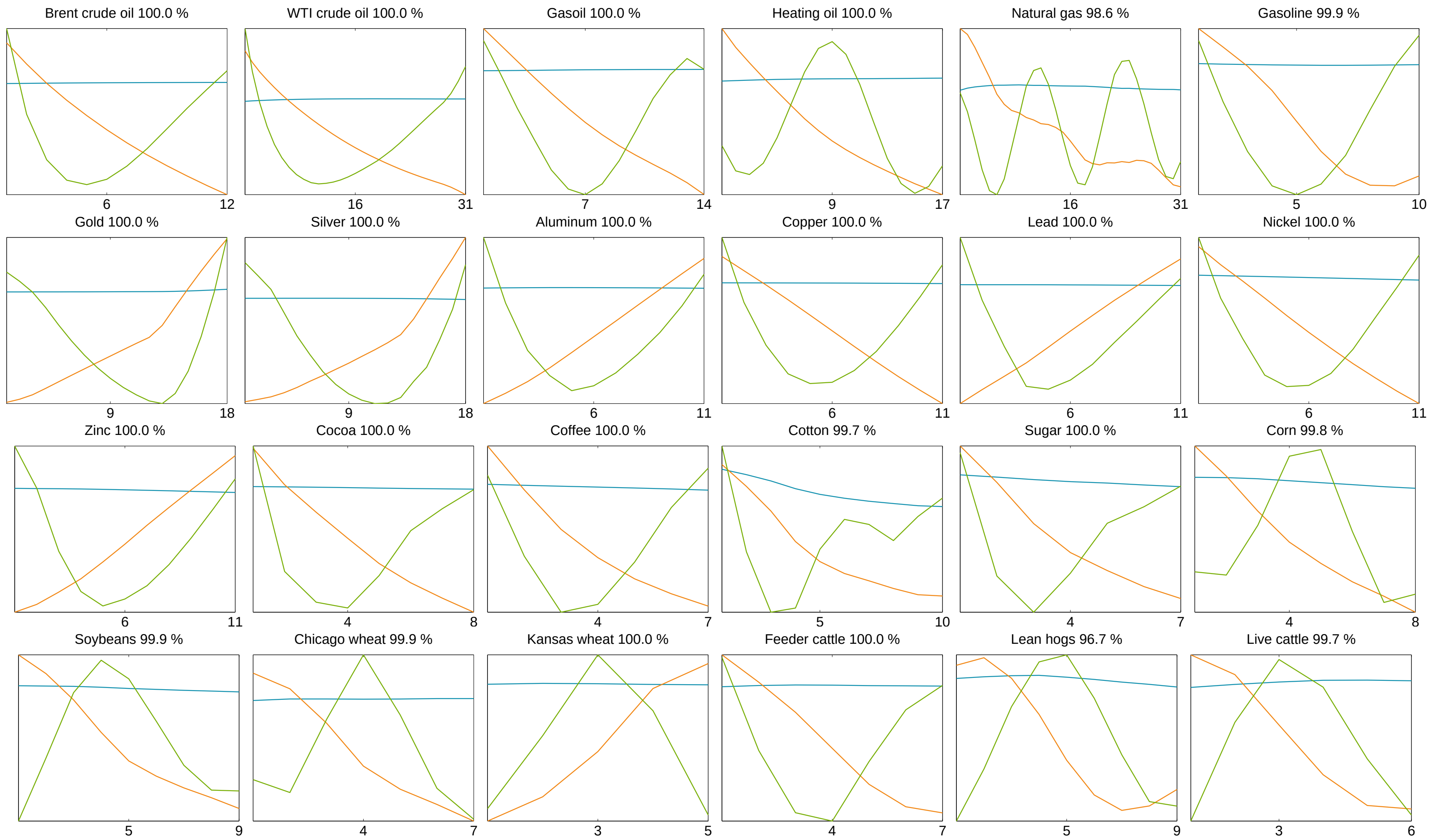
Table 4.C.1 All individual model results

The table presents the estimation results and the fit of the individual commodity state space models. We show the results of three different specification applied to all individual commodities. The two factor (2F) model contains only a level and slope factor; the three factor (3F) model includes a curvature factor; and the 3FS model adds a seasonal correction term. For each model we present the estimated parameter values for the decay parameter λ, and (if relevant) the exposure κ to the seasonal term and the most expensive contract expiry month θ, where $\theta = 0$ corresponds to January. The last column presents the volatility of errors $\sigma(\nu)$. Standard errors of all estimates are provided between brackets.

		two factor (2F)				three factor (3F)				three factor and seasonal (3F)							
Sector	Commodity	λ		$\sigma(\nu)$	R^2	λ		$\sigma(\nu)$	R^2	λ		κ		θ		$\sigma(\nu)$	R^2
Energy	Brent crude oil	0.884	(0.01)	1.28%	99.96%	1.144	(0.01)	0.65%	99.99%	1.143	(0.01)	0.04	(0.99)	1.1	(0.02)	0.65%	99.99%
	WTI crude oil	1.262	(0.01)	1.49%	99.95%	1.272	(0.01)	0.79%	99.99%	1.273	(0.01)	0.03	(1.24)	0.1	(0.03)	0.79%	99.99%
	Gasoil	1.644	(0.02)	1.41%	99.96%	1.991	(0.04)	1.17%	99.97%	1.688	(0.04)	0.92	(0.02)	0.4	(0.00)	1.03%	99.98%
	Heating oil	1.956	(0.04)	2.40%	99.88%	3.392	(0.06)	2.17%	99.90%	3.600	(0.03)	2.80	(0.01)	0.8	(0.00)	1.21%	99.97%
	Natural gas	0.822	(0.03)	6.39%	98.07%	1.171	(0.03)	5.72%	98.45%	1.137	(0.02)	6.38	(0.01)	0.9	(0.00)	3.56%	99.40%
	Gasoline	2.298	(0.07)	3.59%	99.67%	4.346	(0.04)	3.11%	99.75%	3.285	(0.04)	4.78	(0.01)	6.2	(0.00)	1.52%	99.94%
Metals	Gold	0.011	(6.36)	0.44%	99.99%	0.129	(0.08)	0.19%	100.00%	0.129	(0.08)	0.00	(22.65)	0.1	(0.79)	0.19%	100.00%
	Silver	0.026	(1.79)	0.54%	99.99%	0.278	(0.05)	0.35%	100.00%	0.278	(0.05)	0.02	(2.08)	2.6	(0.04)	0.35%	100.00%
	Aluminum	0.187	(0.04)	0.77%	99.91%	0.574	(0.01)	0.32%	99.98%	0.574	(0.01)	0.01	(2.81)	9.6	(0.03)	0.32%	99.98%
	Copper	0.111	(0.10)	1.13%	99.97%	0.489	(0.01)	0.41%	100.00%	0.489	(0.01)	0.02	(1.35)	7.3	(0.02)	0.41%	100.00%
	Lead	0.324	(0.13)	0.57%	99.99%	1.771	(0.02)	0.36%	100.00%	1.767	(0.02)	0.04	(0.72)	9.2	(0.02)	0.36%	100.00%
	Nickel	0.095	(0.28)	0.91%	99.98%	1.250	(0.01)	0.71%	99.99%	1.250	(0.01)	0.00	(30.31)	4.5	(0.17)	0.71%	99.99%
	Zinc	0.059	(0.19)	0.69%	99.97%	1.132	(0.01)	0.32%	99.99%	1.131	(0.01)	0.03	(0.79)	9.6	(0.02)	0.32%	99.99%
Softs	Cocoa	0.360	(0.10)	0.74%	99.96%	1.413	(0.03)	0.58%	99.97%	1.415	(0.03)	0.08	(0.71)	8.8	(0.02)	0.58%	99.97%
	Coffee	0.954	(0.03)	0.80%	99.96%	1.468	(0.03)	0.57%	99.98%	1.468	(0.03)	0.00	(56.48)	0.1	(1.61)	0.57%	99.98%
	Cotton	1.484	(0.03)	1.76%	99.45%	3.566	(0.03)	1.37%	99.67%	3.482	(0.03)	0.91	(0.06)	5.9	(0.01)	1.21%	99.74%
	Sugar	1.610	(0.05)	2.13%	99.76%	3.585	(0.04)	1.53%	99.88%	3.272	(0.04)	1.13	(0.05)	2.5	(0.01)	1.32%	99.91%
Grains	Corn	1.603	(0.03)	2.37%	99.57%	0.696	(0.18)	2.10%	99.67%	2.743	(0.04)	1.80	(0.03)	5.7	(0.00)	1.56%	99.81%
	Soybeans	1.284	(0.05)	1.82%	99.73%	3.536	(0.03)	1.46%	99.83%	3.350	(0.03)	1.30	(0.03)	5.7	(0.00)	1.19%	99.89%
	Chicago wheat	1.345	(0.05)	2.09%	99.70%	1.143	(0.15)	1.86%	99.76%	1.496	(0.17)	1.53	(0.04)	2.0	(0.00)	1.51%	99.84%
	Kansas wheat	1.027	(0.11)	1.94%	99.72%	3.318	(0.07)	1.64%	99.80%	2.461	(0.08)	1.44	(0.05)	2.2	(0.01)	1.35%	99.86%
Meats	Feeder cattle	2.532	(0.07)	0.96%	99.83%	4.522	(0.05)	0.78%	99.89%	5.221	(0.07)	0.73	(0.07)	10.0	(0.01)	0.71%	99.91%
	Lean hogs	1.701	(0.18)	6.66%	89.78%	4.022	(0.04)	5.12%	93.95%	4.109	(0.07)	7.78	(0.02)	6.1	(0.00)	3.09%	97.80%
	Live cattle	3.200	(0.10)	2.19%	98.85%	4.099	(0.05)	1.77%	99.25%	4.487	(0.10)	2.32	(0.03)	1.9	(0.00)	1.29%	99.60%

Figure 4.C.2 Principal components

These figures show for each commodity the first three principal component loadings extracted from the covariance matrix of raw prices. Contracts are included when they have valid observations for at least 90% of the months in our sample.



Preliminary commonality results

A preliminary check for commonality is obtained by comparing the unobserved level, slope and curvature factors from the individual models. Using PCA, we check to what extent the factors can be explained by the first principal component. The market-wide level component is approximated by the first principal component when PCA is applied to all 24 extracted level factors. For each commodity we can compute the fraction of "individual" variance that is explained by this first principal component. We then take the average over all commodities in a particular sector. The same analysis is also applied on all slope factors and on all curvature factors. If the first principal explains a large part of the factors variations, it is an indication for commonality.

Panel A in Table 4.C.2 shows that there seems to be a market-wide component that drives the level factors as the first principal component explains 78.7% of the variation in individual level factors. Especially, the energy, metals, and grains level factors comove with the market-wide level component, as more than 80% of their variation is explained. For the other two commodity sectors we observe that 56.1% and 77.3% of their variation is explained by the market-wide level component. Investigating the market-wide slope component shows that there is less comovement on average indicated by the explained variation of 26.6%. The decomposition in sectors shows that the market-wide slope component still explains half of the variation in the energy slope factors, while it hardly explains variation for the softs and grains sectors. The market-wide curvature component shows similar results as the market-wide slope component. In general, the market-wide curvature component explains 19.9% of the variation in the individual commodity curvature factors. For energy the percentage of explained variation is much higher (38.2%), while for the softs sector it explains none of the variation at all.

Table 4.C.2 PCA commonality results

The table presents the percentage of explained variation of the extracted level, slope and curvature factors for the first three principal components. Panel A shows how much variation is explained by the market-wide level, slope and curvature components. In Panel B the percentages refer to the explained variation when PCA is applied to a particular sector, i.e. this variation can be due to both the market-wide and sector specific component. The differences in percentages explained variation between Panel A and B, indicate the comovement due to the sector specific component.

	level factor			slope factor			curvature factor		
	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
Panel A Market-wide component									
1) all commodities	78.7%	9.8%	3.5%	26.5%	17.0%	14.9%	19.9%	16.7%	9.4%
2) energy	80.9%	16.3%	0.3%	58.6%	2.4%	22.8%	38.2%	5.9%	12.4%
3) metals	86.5%	1.4%	6.7%	23.9%	12.5%	25.8%	-	-	-
4) softs	56.1%	22.4%	2.7%	6.5%	13.2%	1.4%	0.0%	0.0%	0.0%
5) grains	85.5%	7.0%	1.1%	1.9%	61.5%	4.0%	0.1%	1.1%	7.9%
6) meats	77.3%	3.4%	6.7%	27.9%	2.6%	6.1%	6.5%	11.5%	12.1%
Panel B Market-wide and Sector component									
1) energy	93.5%	6.4%	0.0%	79.1%	11.7%	4.9%	46.0%	21.4%	12.4%
2) metals	90.3%	5.8%	1.7%	49.9%	24.4%	12.2%	-	-	-
3) softs	77.3%	12.8%	6.9%	43.4%	26.3%	20.3%	33.6%	28.3%	24.1%
4) grains	97.8%	1.3%	0.8%	69.2%	19.3%	9.4%	60.4%	20.1%	11.0%
5) meats	90.4%	9.0%	0.6%	62.6%	26.2%	11.2%	47.6%	35.3%	17.1%

4.D Additional Schwartz comparison results

This section presents results on the differences in model fit between our Nelson-Siegel framework and the Schwartz (1997) model. In Section 4.4.1 Table 4.2 we present the fit of our individual models. In Table 4.D.1 we present these same numbers together with the R^2 of the corresponding Schwartz (1997) model. For all commodities our model has a better fit. The minimum fit increase is 1.6%, the average increase is 5.6% and the maximum increase is 29.2%.

Table 4.D.1 Model fit comparison

This table compares the differences in fit $\left(R^2\right)$ between the Schwartz (1997) model and our Nelson-Siegel type model. The difference is computed as $(R^2_{\text{Nelson-Siegel}} - R^2_{\text{Schwartz}})/R^2_{\text{Nelson-Siegel}}$.

Commodity	Schwartz	Nelson-Siegel	diff.	Commodity	Schwartz	Nelson-Siegel	diff
Brent crude oil	97.06%	99.99%	3.0%	Cocoa	93.22%	99.97%	7.2%
WTI crude oil	97.28%	99.99%	2.8%	Coffee	94.76%	99.98%	5.5%
Gasoil	98.40%	99.98%	1.6%	Cotton	88.66%	99.74%	12.5%
Heating oil	98.41%	99.97%	1.6%	Sugar	96.51%	99.91%	3.5%
Natural gas	92.14%	99.40%	7.9%				
Gasoline	97.43%	99.94%	2.6%	Corn	95.72%	99.81%	4.3%
				Soybeans	95.32%	99.89%	4.8%
Gold	96.73%	99.99%	3.4%	Chicago wheat	95.47%	99.84%	4.6%
Silver	96.13%	99.99%	4.0%	Kansas wheat	95.15%	99.86%	5.0%
Aluminum	90.29%	99.91%	10.7%				
Copper	96.99%	99.97%	3.1%	Feeder cattle	97.51%	99.91%	2.5%
Lead	98.05%	99.99%	2.0%	Lean hogs	75.68%	97.80%	29.2%
Nickel	96.45%	99.98%	3.7%	Live cattle	95.77%	99.60%	4.0%
Zinc	94.80%	99.97%	5.5%				

4.E Macroeconomic data

Table 4.E.1 lists all macroeconomic and commodity specific data series. For each series we provide their name, code, transformation, category, source, and description. An overview of the categories is shown in Table 4.E.2 while the transformation codes are explained in Table 4.E.3.

Table 4.E.1 Data series

The table presents all macroeconomic and commodity specific data series (in line with Stock and Watson, 2012 and Gorton, Hayashi, and Rouwenhorst, 2013). For each series we provide their name, code, transformation, category, source, and description. An overview of the categories is shown in Table 4.E.2 while the transformation codes are explained in Table 4.E.3. The codes correspond to the database identifiers of the source.

Used abbreviations: St. Louis, Federal Reserve Economic Data (FRED); Commodity Futures Trading Commission (CFTC); Department of Energy (DOE); Intercontinental Exchange (ICE); U.S. Department of Agriculture (USDA).

[a] Author's website:

[b] We follow the details given in Appendix B of Gorton, Hayashi, and Rouwenhorst (2013)

Name	Code	T	Cat	Source	Description
Cons-Dur	DNDGRG3M086SBEA	5	1	FRED	Personal consumption expenditures: Nondurable goods, Price index (2009=100), SA
Cons-NonDur	DPCERA3M086SBEA	5	1	FRED	Real personal consumption expenditures, Quantity index (2009=100), SA
Cons-Serv	DSERRG3M086SBEA	5	1	FRED	Personal consumption expenditures: Services, Price index (2009=100), SA
Exports	USEXNGS.B	5	1	Datastream	Real exports
Imports	USIMNGS.B	5	1	Datastream	Real imports
IP: cons dble	IPDCONGD	5	2	FRED	Industrial Production: Durable Consumer Goods Index (2007=100), SA
IP: cons nondble	IPNCONGD	5	2	FRED	Industrial Production: Nondurable Consumer Goods Index (2007=100), SA
IP: bus eqpt	IPBUSEQ	5	2	FRED	Industrial Production: Business Equipment Index (2007=100), SA
IP: dble mats	IPDMAT	5	2	FRED	Industrial Production: Durable Materials Index (2007=100), SA
IP: nondble mats	IPNMAT	5	2	FRED	Industrial Production: nondurable Materials Index (2007=100), SA
IP: mfg	IPMANSICS	5	2	FRED	Industrial Production: Manufacturing (SIC) Index (2007=100), SA
IP: fuels	IPUTIL	5	2	FRED	Industrial Production: Electric and Gas Utilities Index (2007=100), SA
NAPM prodn	NAPMPI	1	2	FRED	ISM Manufacturing: Production Index, SA
Capacity Util	TCU	1	2	FRED	Capacity Utilization: Total Industry % of Capacity, SA

Continued on next page

Name	Code	T	Cat	Source	Description
Emp: mining	CES1021000001	5	3	FRED	All Employees: Mining and Logging: Mining, Thous. of Persons, SA
Emp: const	USCONS	5	3	FRED	All Employees: Construction, Thous. of Persons, SA
Emp: dble gds	DMANEMP	5	3	FRED	All Employees: Durable goods, Thous. of Persons, SA
Emp: nondbles	NDMANEMP	5	3	FRED	All Employees: Nondurable goods, Thous. of Persons, SA
Emp: services	SRVPRD	5	3	FRED	All Employees: Service-Providing Industries, Thous. of Persons, SA
Emp: TTU	USTPU	5	3	FRED	All Employees: Trade, Transportation & Utilities, Thous. of Persons, SA
Emp: wholesale	USWTRADE	5	3	FRED	All Employees: Wholesale Trade, Thous. of Persons, SA
Emp: retail	USTRADE	5	3	FRED	All Employees: Retail Trade, Thous. of Persons, SA
Emp: FIRE	USFIRE	5	3	FRED	All Employees: Financial Activities, Thous. of Persons, SA
Emp: Govt	USGOVT	5	3	FRED	All Employees: Government, Thous. of Persons, SA
Emp. Hours	AWHI	5	3	FRED	Aggr. Wkly Hours: Prod. and Nonsuperv. Employ.: Total Private Industries (2002=100), SA
Avg hrs	CES0600000007	1	3	FRED	Avg. Wkly Hours of Prod. and Nonsuperv. Employ.: Goods-Producing Hours, SA
Overtime: mfg	AWOTMAN	2	3	FRED	Avg. Wkly Overtime Hours of Prod. and Nonsuperv. Employees: Manufacturing Hours, SA
U: all	UNRATE	2	4	FRED	Unemployment rate: all workers, 16 years and over, Percentage, SA
U: mean duration	UEMPMEAN	2	4	FRED	Average (Mean) Duration of Unemployment, Weeks, SA
U: < 5 wks	UEMPLT5	5	4	FRED	Number of Civilians Unemployed - Less Than 5 Weeks, Thous. of Persons, SA
U: 5-14 wks	UEMP5TO14	5	4	FRED	Number of Civilians Unemployed for 5 to 14 Weeks, Thous. of Persons, SA
U: 15+ wks	UEMP15OV	5	4	FRED	Number of Civilians Unemployed for 15 Weeks and Over, Thous. of Persons, SA
U: 15-26 wks	UEMP15T26	5	4	FRED	Number of Civilians Unemployed for 15 to 26 Weeks, Thous. of Persons, SA
U: 27+ wks	UEMP27OV	5	4	FRED	Number of Civilians Unemployed for 27 Weeks and Over, Thous. of Persons, SA
HStarts: NE	HOUSTNE	4	5	FRED	Housing Starts in Northeast Census Region, Thous. of Units, SAAR
HStarts: MW	HOUSTMW	4	5	FRED	Housing Starts in Midwest Census Region, Thous. of Units, SAAR
HStarts: S	HOUSTS	4	5	FRED	Housing Starts in South Census Region, Thous. of Units, SAAR
HStarts: W	HOUSTW	4	5	FRED	Housing Starts in West Census Region, Thous. of Units, SAAR
PMI	NAPM	1	6	FRED	ISM Manufacturing: PMI Composite Index, SA
NAPM new orders	NAPMNOI	1	6	FRED	ISM Manufacturing: New Orders Index, SA
NAPM vendor del	NAPMSDI	1	6	FRED	ISM Manufacturing: Supplier Deliveries Index, SA

Continued on next page

Name	Code	T	Cat	Source	Description
NAPM Invent	NAPMII	1	6	FRED	ISM Manufacturing: Inventories Index, NSA
Orders (ConsGoods)	ACOGNO	5	6	FRED	Manufacturers New Orders for Cons. Goods Indus., Mil. of $, SA
Orders (NDCapGoods)	ANDENO	5	6	FRED	Manufacturers New Orders for Capital Goods: Nondef. Capital Goods Indus., Mil. of $, SA
CPI-core	CPIULFSL	6	7	FRED	Consumer Price Index for All Urban Cons.: All Items Less Food, Index (1982-84=100), SA
PCED	PCEPI	6	7	FRED	Personal Consumption Expenditures, Price Index (2009=100), SA
AHE: const	CES2000000008	5	8	FRED	Avg. Hourly Earnings of Prod. and Nonsuperv. Employees: Construction, $ per Hour, SA
AHE: mfg	CES3000000008	5	8	FRED	Avg. Hourly Earnings of Prod. and Nonsuperv. Employees: Manufacturing, $ per Hour, SA
FedFunds	FEDFUNDS	2	9	FRED	Effective Federal Funds Rate, % per annum, NSA
3mo T-bill	TB3MS	2	9	FRED	3-Month Treasury Bill: Secondary Market Rate, % per annum, NSA
3mo T-bill	TB6MS	2	9	FRED	6-Month Treasury Bill: Secondary Market Rate, % per annum, NSA
M1	M1SL	6	10	FRED	M1 Money Stock, Bil. of $, SA
M2	M2SL	6	10	FRED	M2 Money Stock, Bil. of $, SA
MB	AMBSL	6	10	FRED	St. Louis Adjusted Monetary Base, Bil. of $, SA
Reserves tot.	TOTRESNS	6	10	FRED	Total Reserves of Depository Institutions, Bil. of $, NSA
BUSLOANS	BUSLOANS	6	10	FRED	Commercial and Industrial Loans, All Commercial Banks, Bil. of $, SA
Cons credit	NONREVSL	6	10	FRED	Total Nonrevolving Credit Owned and Securitized, Outstanding, Bil. of $, SA
Ex rate: avg	TWEXMMTH	5	11	FRED	Trade Weighted U.S. Dollar Index: Major Currencies, Index (Mar 1973=100), NSA
Ex rate: Switz	EXSZUS	5	11	FRED	Switzerland / U.S. Foreign Exchange Rate Swiss Francs to 1 U.S., $, NSA
Ex rate: Japan	EXJPUS	5	11	FRED	Japan / U.S. Foreign Exchange Rate Japanese Yen to 1 U.S., $, NSA
Ex rate: UK	EXUSUK	5	11	FRED	U.S. / U.K. Foreign Exchange Rate U.S. $ to 1 British Pound, £, NSA
Ex rate: Canada	EXCAUS	5	11	FRED	Canada / U.S. Foreign Exchange Rate Canadian $ to 1 U.S., $, NSA
S&P 500	SP500	5	12	FRED	S&P 500, Index, NSA
DJIA	USSHRPRCF	5	12	Datastream	Dow Jones Industrial Average
Consumer expect	UMCSENT	2	13	FRED	University of Michigan: Consumer Sentiment, Index (1966Q1=100), NSA
ADS	ADS	1	14	FRB of Phil.	Aruoba Diebold Scotti financial conditions index

Continued on next page

Name	Code	T	Cat	Source	Description
Sentiment	Sentiment	1	14	Author's website[a]	Baker Wurgler paper
Baltic Dry index	BALTICF	1	14	Datastream	Dry bulk shipping price
Comm. vol.	CRBSPOT	1	15	Datastream	Volatility of Commodity Research Bureau (CRB) spot market price index
HP WTI	hedging WTI	1	16A	CFTC	Hedging pressure computed from Commitment of Traders Reports
HP gasoline	hedging GL	1	16A	CFTC	Hedging pressure computed from Commitment of Traders Reports
HP heating oil	hedging HO	1	16A	CFTC	Hedging pressure computed from Commitment of Traders Reports
HP natural gas	hedging NG	1	16A	CFTC	Hedging pressure computed from Commitment of Traders Reports
HP copper	hedging CP	1	16B	CFTC	Hedging pressure computed from Commitment of Traders Reports
HP cocoa	hedging CC	1	16C	CFTC	Hedging pressure computed from Commitment of Traders Reports
HP coffee	hedging CF	1	16C	CFTC	Hedging pressure computed from Commitment of Traders Reports
HP sugar	hedging SG	1	16C	CFTC	Hedging pressure computed from Commitment of Traders Reports
HP cotton	hedging CT	1	16C	CFTC	Hedging pressure computed from Commitment of Traders Reports
HP soybeans	hedging S	1	16D	CFTC	Hedging pressure computed from Commitment of Traders Reports
HP wheat	hedging W	1	16D	CFTC	Hedging pressure computed from Commitment of Traders Reports
HP corn	hedging C	1	16D	CFTC	Hedging pressure computed from Commitment of Traders Reports
HP feeder cattle	hedging FC	1	16E	CFTC	Hedging pressure computed from Commitment of Traders Reports
HP live cattle	hedging LC	1	16E	CFTC	Hedging pressure computed from Commitment of Traders Reports
HP lean hogs	hedging LH	1	16E	CFTC	Hedging pressure computed from Commitment of Traders Reports
inventory crude oil	inventory WTI	5	17A	DOE	U.S. ending stocks excluding SPR of crude oil, thousands of barrels[b]
inventory gasoline	inventory GL	5	17A	DOE	U.S. motor gasoline ending stocks, thousands of barrels[b]
inventory heating oil	inventory HO	5	17A	DOE	U.S. total stocks of distillate fuel oil[b]
inventory natural gas	inventory NG	5	17A	DOE	U.S. total natural gas in underground storage (working gas), millions of cubic feet[b]
inventory gold	COMXGOLD Index	5	17B	Bloomberg	Comex warehouse stocks[b]
inventory silver	COMXSILV Index	5	17B	Bloomberg	Comex warehouse stocks[b]
inventory copper	LSCA Index	5	17B	Bloomberg	LME warehouse stocks[b]
inventory aluminum	LSAH Index	5	17B	Bloomberg	LME warehouse stocks[b]
inventory lead	LSPB Index	5	17B	Bloomberg	LME warehouse stocks[b]

Continued on next page

Name	Code	T	Cat	Source	Description
inventory nickel	LSNI Index	5	17B	Bloomberg	LME warehouse stocks[b]
inventory zinc	LSZS Index	5	17B	Bloomberg	LME warehouse stocks[b]
inventory cocoa	inventory CC	5	17C	ICE	ICE wareh. stocks (ports of New York, Delaware River, Hampton Roads, Albany, Baltimore)[b]
inventory coffee	inventory CF	5	17C	ICE	ICE wareh. stocks (ports of New York, New Orleans, Houston, Miami, Antwerp, Hamburg, Barcelona)[b]
inventory cotton	inventory CT	5	17C	ICE	Hist. Certif. Stock Report (ports of Dallas, Galveston, Greenville, Houston, Memphis, New Orleans)[b]
inventory sugar	inventory SG	5	17C	USDA	U.S. sugar stocks held by primary distributors[b]
inventory corn	inventory C	5	17D	USDA	Stocks of Grain at Selected Terminals and Elevator Sites, thousands of bushels[b]
inventory soybeans	inventory S	5	17D	USDA	Stocks of Grain at Selected Terminals and Elevator Sites, Thousands of Bushels[b]
inventory wheat	inventory W	5	17D	USDA	Stocks of Grain at Selected Terminals and Elevator Sites, Thousands of Bushels[b]
inventory feeder cattle	inventory FC	5	17E	USDA	United States Cattle Placed on Feed in 7 States[b]
inventory live cattle	inventory LC	5	17E	USDA	Frozen beef stocks in cold storage in the U.S.[b]
inventory lean hogs	inventory LH	5	17E	USDA	Frozen pork stocks in cold storage in the U.S.[b]

Table 4.E.2 Data categories
The table presents the data category codes (following Stock and Watson, 2012).

Category code	Category name
1	GDP components
2	Industrial production
3	Employment
4	Unemployment rate
5	Housing
6	Business inventories
7	Prices
8	Wages
9	Interest rates
10	Money
11	Exchange rates
12	Stock prices
13	Consumer expectations
14	Financial conditions
15	Commodity volatility
16	Hedging pressure
17	Commodity inventories

Table 4.E.3 Data transformations
The table presents the data transformation codes (following Stock and Watson, 2012). Z_t denotes the raw series and X_t the transformed series used to compute the principal components.

Transformation code	X_t
1	Z_t
2	$Z_t - Z_{t-1}$
3	$(Z_t - Z_{t-1}) - (Z_{t-1} - Z_{t-2})$
4	$\ln(Z_t)$
5	$\ln(Z_t/Z_{t-1})$
6	$\ln(Z_t/Z_{t-1}) - \ln(Z_{t-1}/Z_{t-2})$

Nederlandse samenvatting (Summary in Dutch)

Sta open voor nieuwe ideeen

Abbott (1884)

Een inspirerend en intrigerend boek over dimensies is "Flatland - A Romancy of Many Dimensions" van Edwin A. Abbott.[22] In deze satirische roman beschrijft Abbott werelden met verschillende dimensies vanuit het oogpunt van Een Vierkant. E. (Een) Vierkant leeft in de twee-dimensionele wereld genaamd *Flatland*. E. Vierkant bezoekt in een droom de één-dimensionale wereld (*Lineland*) bewoond door Punten. Hij probeert hun tevergeefs te overtuigen van het bestaan van een tweede dimensie. Wanneer E. Vierkant wordt bezocht door een Bol (afkomstig uit *Spaceland*), kan hij niet bevatten dat een derde dimensie bestaat totdat hij *Spaceland* zelf bezoekt. Nu E. Vierkant openstaat voor nieuwe dimensies, probeert hij de Bol te overtuigen van de theoretische kans dat er een vierde (en vijfde, zesde, ...) dimensie bestaat. Hij roept de woede van de Bol over zich af en wordt terug gebracht naar zijn twee-dimensionale wereld. E. Vierkant herkent de ontkenning (van de Bol en Punten) van het bestaan van hogere dimensies omdat hijzelf in eerste instantie ook niet te overtuigen was.

De link tussen *Flatland* en dit proefschrift is het concept van dimensies. De financiële wereld bestaat uit veel meer dimensies dan wij ons voor kunnen stellen als bewoners van onze drie-dimensionale wereld. De twee die in dit proefschrift centraal staan zijn de dimensies van liquiditeit en de termijnstructuur dimensie in de grondstoffen markt. Beiden zijn nog niet volledig onderzocht. Naast de directe link is er ook een andere link tussen het boek,

[22]Hoewel Flatland niet genegeerd werd op het moment van publicatie, was het geen groot succes. Het boek werd opnieuw ontdekt nadat Albert Einsteins relativiteitstheorie werd gepubliceerd en het werd genoemd in *Nature* (Garnett, 1920).

Flatland, en dit proefschrift. De onderliggende boodschap van het boek is dat we "open moeten staan voor nieuwe ideeën", wat vergelijkbaar is met het PhD-traject. We moeten nieuwe gebieden en toepassingen onderzoeken en we moeten ontvankelijk zijn voor zaken die we ons in eerste instantie niet kunnen voorstellen.

Het doel van dit proefschrift is om raamwerken te presenteren die een beter beeld geven van de eerder genoemde dimensies. De nieuwe resultaten en inzichten die hieruit volgen zijn relevant voor investeerders en academici. Dit proefschrift bestaat uit twee delen. In het eerste deel staat liquiditeit centraal terwijl het tweede deel de focus legt op de termijnstructuur in grondstofmarkten.

Hoofdstuk 2 is gebaseerd op Karstanje, Sojli, Tham en van der Wel (2013). In dit hoofdstuk focussen we op de dimensies van liquiditeit voor het gebruik van markt *timing*, het op het juiste moment in- of uit de markt stappen. We vergelijken de prestaties van dynamische asset allocatie strategieën, die allen gedreven worden door de korte termijn voorspelkracht van maandelijkse aandelen rendementen door één van de liquiditeitsmaatstaven. We vinden drie hoofdresultaten: markt *timing* op basis van liquiditeit leidt tot tastbare economische winsten; een risico-averse investeerder is bereid om een hogere vergoeding te betalen om te wisselen naar de *Zeros* liquiditeitsmaatstaf (Lesmond, Ogden en Trzcinka, 1999); de *Zeros* liquiditeitsmaatstaf presteert beter dan de andere maatstaven door zijn robuustheid tegen extreme financiële situaties. Deze bevindingen zijn onafhankelijk van de bekeken tijdsperiode en zijn robuust tegen het controleren voor bestaande marktrendement voorspellers of het gebruik van voor risico gecorrigeerde rendementen.

Hoofdstuk 3 is gebaseerd op de Groot, Karstanje en Zhou (2014). We onderzoeken nieuwe "momentum" strategieën in grondstof termijncontracten (zogenaamde *futures*), waarbij informatie van de termijnstructuur wordt meegenomen. De termijnstructuur, of grondstof curve, is de verzameling van termijncontracten met verschillende looptijden maar dezelfde onderliggende grondstof. We laten zien dat momentum strategieën die investeren in contracten op de curve waar de verwachte "rol-rendement" het grootst is of waar momentum het best heeft gepresteerd, significant hogere rendementen behalen dan een standaard momentum strategie. De standaard strategie handelt enkel in de eerst aflopende termijncontracten. Wanneer we conservatieve transactiekosten meenemen in onze analyse, observeren we dat onze lage-"turnover" momentum strategie het netto rendement meer dan verdubbelt ten opzichte van een traditionele momentum strategie.

Hoofdstuk 4 is gebaseerd op Karstanje, van der Wel en van Dijk (2014). In dit hoofdstuk onderzoeken we de samenhang tussen de drijvende factoren achter grondstof curven. We gebruiken het raamwerk van het dynamische Nelson-Siegel (1987) model waardoor we zowel de samenhang in prijsniveaus als in termijnstructuur vormen (gekarakteriseerd door hun helling en kromming) kunnen onderzoeken. Onze empirische resultaten, gebaseerd op data van 24 grondstoffen over de periode 1995-2012, laten zien dat individuele grondstof curven worden gedreven door gemeenschappelijke componenten. Het gemeenschappelijk deel is vooral sector specifiek, wat impliceert dat grondstoffen een heterogene asset klasse zijn. Het gemeenschappelijk deel van het niveau van de curve is groter geworden over de tijd. Deze toename valt tegelijkertijd met de financialisering van de grondstofmarkt. De marktbrede niveau component, die alle grondstoffen beïnvloedt, is gerelateerd aan economische output variabelen, wisselkoersen en *hedging pressure* (het relatieve verschil tussen het aantal speculanten en producenten). De drijvende factoren achter de vorm van de curve zijn gerelateerd aan grondstofvoorraden (*theory of storage*), *hedging pressure* (*theory of normal backwardation*) en de rentetermijnstructuur. Het gebruik van volledige curve data verandert de resultaten met betrekking tot de onderlinge grondstof samenhang, vergeleken met het gebruik van data van enkel eerst aflopende termijncontracten. De resultaten geven meer inzicht in de dynamiek van de grondstofmarkten en kunnen helpen bij de constructie van grondstof investeringsportfolios en *hedging* beslissingen.

Bibliography

Abbott, E. A., 1884. Flatland: A Romance of Many Dimensions. Seely & Co.

Acharya, V. V., Pedersen, L. H., 2005. Asset pricing with liquidity risk. Journal of Financial Economics 77 (2), 375 – 410.

Ai, C., Chatrath, A., Song, F., 2006. On the comovement of commodity prices. American Journal of Agricultural Economics 88 (3), 574–588.

Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets 5 (1), 31 – 56.

Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid-ask spread. Journal of Financial Economics 17 (2), 223 – 249.

Amihud, Y., Mendelson, H., Lauterbach, B., 1997. Market microstructure and securities values: Evidence from the Tel Aviv Stock Exchange. Journal of Financial Economics 45 (3), 365 – 390.

Amihud, Y., Mendelson, H., Pedersen, L. H., 2005. Liquidity and asset prices. Foundations and Trends in Finance 1 (4), 269–364.

Aruoba, S., Diebold, F., Scotti, C., 2009. Real-time measurement of business conditions. Journal of Business and Economic Statistics 27 (4), 417–427.

Asness, C., Moskowitz, T., Pedersen, L., 2009. UPDATEvalue and momentum everywhere. Working paper, AFA 2010 Atlanta Meetings Paper.

Asparouhova, E., Bessembinder, H., Kalcheva, I., 2010. Liquidity biases in asset pricing tests. Journal of Financial Economics 96 (2), 215–237.

Asparouhova, E., Bessembinder, H., Kalcheva, I., 2013. Noisy prices and inference regarding returns. Journal of Finance 68 (2), 665–714.

Baker, M., Stein, J. C., 2004. Market liquidity as a sentiment indicator. Journal of Financial Markets 7 (3), 271 – 299.

Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. Journal of Finance 61 (4), 1645–1680.

Bakshi, G., Panayotov, G., Skoulakis, G., 2011. The baltic dry index as a predictor of global stock returns, commodity returns, and global economic activity. Working paper.

Baltas, A.-N., Kosowski, R., 2013. Momentum strategies in futures markets and trend-following funds. Tech. rep.

Bekaert, G., Harvey, C. R., Lundblad, C., 2007. Liquidity and expected returns: Lessons from emerging markets. Review of Financial Studies 20 (6), 1783 – 1831.

Ben-Rephael, A., Kadan, O., Wohl, A., 2010. The diminishing liquidity premium. Working paper.

Benston, G., Hagerman, R., 1974. Determinants of bid-ask spreads in the over-the-counter market. Journal of financial economics 1 (4), 353–364.

Berkman, H., Eleswarapu, V. R., 1998. Short-term traders and liquidity: A test using Bombay Stock Exchange data. Journal of Financial Economics 47 (3), 339 – 355.

Bessembinder, H., 1992. Systematic risk, hedging pressure, and risk premiums in futures markets. Review of Financial Studies 5 (4), 637–667.

Brennan, M., 1991. The price of convenience and the valuation of commodity contingent claims. In: Lund, D., Oksendal, B. (Eds.), Stochastic Models and Option Values: Applications to Resources, Environment and Investment Problems. North-Holland, New York.

Brennan, M. J., 1958. The supply of storage. American Economic Review 48, 50–72.

Brennan, M. J., Chordia, T., Subrahmanyam, A., 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. Journal of Financial Economics 49 (3), 345 – 373.

Brennan, M. J., Schwartz, E., 1985. Evaluating natural resource investments. Journal of Business 58, 135–157.

Brooks, C., Prokopczuk, M., 2013. The dynamics of commodity prices. Quantitative Finance 13 (4), 527–542.

Byrne, J. P., Fazio, G., Fiess, N., 2012. Primary commodity prices: Co-movements, common factors and fundamentals. Journal of Development Economics 101, 16–26.

Campbell, J. Y., Thompson, S. B., 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? Review of Financial Studies 21 (4), 1509 – 1531.

Cao, C., Chen, Y., Liang, B., Lo, A. W., 2013. Can hedge funds time market liquidity? Journal of Financial Economics 109 (2), 493–516.

Carhart, M. M., 1997. On persistence of mutual fund performance. Journal of Finance 52 (1), 57 – 82.

Carter, C. A., Rausser, G. C., Schmitz, A., 1983. Efficient asset portfolios and the theory of normal backwardation. Journal of Political Economy 91 (2), 319–331.

Casassus, J., Collin-Dufresne, P., 2005. Stochastic convenience yield implied from commodity futures and interest rates. Journal of Finance 60 (5), 2283–2331.

Casassus, J., Liu, P., Tang, K., 2013. Economic linkages, relative scarcity, and commodity futures returns. Review of Financial Studies 26 (5), 1324–1362.

Chang, E. C., 1985. Returns to speculators and the theory of normal backwardation. Journal of Finance 40 (1), 193–208.

Chen, Y.-C., Rogoff, K. S., Rossi, B., 2010. Can exchange rates forecast commodity prices? The Quarterly Journal of Economics 125 (3), 1145–1194.

Cheng, I.-H., Xiong, W., 2014. The financialization of commodities markets. Annual Review of Financial Economics 6 (1), 1–23.

Chordia, T., Huh, S.-W., Subrahmanyam, A., 2009. Theory-based illiquidity and asset pricing. Review of Financial Studies 22 (9), 3629–3668.

Chordia, T., Roll, R., Subrahmanyam, A., 2000. Commonality in liquidity. Journal of Financial Economics 56 (1), 3 – 28.

Chordia, T., Roll, R., Subrahmanyam, A., 2001a. Market liquidity and tradign activity. Journal of Finance 56 (2), 501–350.

Chordia, T., Subrahmanyam, A., Anshuman, R., 2001b. Trading activity and expected stock returns. Journal of Financial Economics 59 (1), 3 – 32.

Chrisie, W., Schultz, P., 1994. Why do NASDAQ market makers avoid odd-eight quotes? Journal of Finance 49 (5), 1813–1840.

Christoffersen, P., Lunde, A., Olesen, K. V., 2014. Factor structure in commodity futures return and volatility. Working paper.

Clements, M., Franses, P., Swanson, N., 2004. Forecasting economic and financial time-series with non-linear models. International Journal of Forecasting 20, 169–183.

Cootner, P., 1960. Returns to speculators: Telser vs. Keynes. Journal of Political Economy 68 (4), 396–404.

Cootner, P., 1967. Speculation and hedging. Tech. rep.

Cortazar, G., Milla, C., Severino, F., 2008. A multicommodity model of futures prices: Using futures prices of one commodity to estimate the stochastic process of another. Journal of Futures Markets 28 (6), 537–560.

Cortazar, G., Naranjo, L., 2006. An N-factor gaussian model of oil futures. The Journal of Futures Markets 26, 243–268.

Corwin, S. A., Schultz, P., 2012. A simple way to estimate bid-ask spreads from daily high and low prices. Journal of Finance 67 (2), 719–760.

Daal, E., Farhat, J., Wei, P. P., 2006. Does futures exhibit maturity effect? new evidence from an extensive set of us and foreign futures contracts. Review of Financial Economics 15 (2), 113–128.

Daskalaki, C., Kostakis, A., Skiadopoulos, G. S., 2014. Are there common factors in individual commodity futures returns? Journal of Banking and Finance 40 (3), 346–363.

de Groot, W., Karstanje, D., Zhou, W., 2014. Exploiting commodity momentum along the futures curves. Journal of Banking and Finance 48 (10), 79–93.

de Roon, F. A., Nijman, T. E., Veld, C., 2000. Hedging pressure effects in futures markets. Journal of Finance 55 (3), 1437–1456.

Deaton, A., Laroque, G., 1992. On the behaviour of commodity prices. Review of Economic Studies 59, 1–23.

Deb, P., Trivedi, P., Varangis, P., 1996. The excess comovement of commodity prices reconsidered. Journal of Applied Econometrics 11, 275–291.

Della-Corte, P., Sarno, L., Thornton, D. L., 2008. The expectation hypothesis of the term structure of very short-term rates: Statistical tests and economic value. Journal of Financial Economics 89 (1), 158–174.

Della-Corte, P., Sarno, L., Tsiakas, I., 2009. An economic evaluation of empirical exchange rate models. Review of Financial Studies 22 (9), 3491–3530.

DeMiguel, V., Garlappi, L., Nogales, F. J., Uppal, R., 2009a. A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. Management Science 55 (5), 798–812.

DeMiguel, V., Garlappi, L., Uppal, R., 2009b. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? Review of Financial Studies 22 (5), 1915 – 1953.

Diebold, F. X., Li, C., 2006. Forecasting the Term Structure of Government Bond Yields. Journal of Econometrics 130, 337–364.

Diebold, F. X., Li, C., Yue, V. Z., 2008. Global yield curve dynamics and interactions: A dynamic Nelson-Siegel approach. Journal of Econometrics 146, 351–363.

Diebold, F. X., Rudebusch, G. D., Aruoba, S. B., Mar. 2006. The macroeconomy and the yield curve: A dynamic latent factor approach. Journal of Econometrics 131 (1-2), 309–338.

Duffee, G., 2011. Forecasting with the term structure: The role of no- arbitrage restrictions. working paper.

Eling, M., Schuhmacher, F., 2007. Does the choice of performance measure influence the evaluation of hedge funds?

Erb, C. B., Harvey, C. R., 2006. The strategic and tactical value of commodity futures. Financial Analysts Journal 62 (2), 69–97.

Fama, E., French, K., 1988. Business cycles and the behavior of metals prices. The Journal of Finance 43 (5), 1075–1093.

Fama, E. F., French, K. R., 1987. Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage. Journal of Business 60 (1), 55–73.

Fama, E. F., French, K. R., 1992. The cross-section of expected stock returns. Journal of Finance 47 (2), 427 – 465.

Fama, E. F., French, K. R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33 (1), 3 – 56.

Fleming, J., Kirby, C., Ostdiek, B., 2001. The economic value of volatility timing. Journal of Finance 56 (1), 329–352.

Fritsch, F., Carlson, R., 1980. Monotone piecewise cubic interpolation. SIAM Journal on Numerical Analysis 17 (2), 238–246.

Fuertes, A.-M., Miffre, J., Rallis, G., 2010. Tactical allocation in commodity futures markets: Combining momentum and term structure signals. Journal of Banking & Finance 34 (10), 2530–2548.

Fung, W., Hsieh, D., 2004. Hedge fund benchmarks; a risk-based approach. Financial Analysts Journal 60 (5), 65–80.

Garber, P. M., 1989. Tulipmania. Journal of Political Economy 97 (3), pp. 535–560.

Garnett, W., 1920. Euclid, newton and einstein. Nature 104 (2624), 627–630.

Geman, H., 2005. Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals and Energy. Wiley Finance.

Geman, H., Nhuyen, V.-N., 2005. Soybean inventory and forward curve dynamics. Management Science 51 (7), 1076–1091.

Gibson, R., Schwartz, E. S., 1990. Stochastic convenience yield and the pricing of oil contingent claims. Journal of Finance 45 (3), 959–976.

Goetzmann, W., Ingersoll, J., Spiegel, M., Welch, I., 2007. Portfolio performance manipulation and manipulation-proof performance measures. The Review of Financial Studies 20 (5), 1503 – 1546.

Gorton, G., Hayashi, F., Rouwenhorst, K., 2013. The fundamentals of commodity futures returns. Review of Finance 17 (1), 35–105.

Gorton, G., Rouwenhorst, K., 2006. Facts and fantasies about commodity futures. Financial Analysts Journal 62 (2), 47–68.

Goyenko, R. Y., Holden, C. W., Trzcinka, C. A., 2009. Do liquidity measures measure liquidity? Journal of Financial Economics 92 (2), 153–181.

Gregoriou, G., Gueyie, J., 2003. Risk-adjusted performance of funds of hedge funds using a modified sharpe ratio. Journal of Wealth Management 6, 77–83.

Hamilton, J. D., Wu, J. C., 2014. Risk premia in crude oil futures. Journal of International Money and Finance 42 (4), 9–37.

Han, Y., 2006. Asset allocation with a high dimensional latent factor stochastic volatility model. Review of Financial Studies 19 (1), 237–271.

Hansen, N. S., Lunde, A., 2013. Analyzing oil futures with a dynamic Nelson-Siegel model. CREATES research paper 2013-36.

Harris, L. E., 1991. Stock price clustering and discreteness. Review of Financial Studies 4 (3), 389–415.

Harris, L. E., 2003. Trading and Exchanges. Oxford University Press, New York.

Hasbrouck, J., 2009. Trading costs and returns for U.S. equities: Estimating effective costs from daily data. Journal of Finance 64 (3), 1445–1477.

Holden, C. W., 2009. New low-frequency spread measures. Journal of Financial Markets 12 (4), 778 – 813.

Jagannathan, R., Ma, T., 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. The Journal of Finance 58 (4), 1651–1683.

Jones, C. M., 2002. A century of stock market liquidity and trading costs. Working paper.

Kaldor, N., 1939. Speculation and economic stability. Review of Economic Studies 7 (1), 1–27.

Karstanje, D., Sojli, E., Tham, W. W., van der Wel, M., 2013. Economic valuation of liquidity timing. Journal of Banking and Finance 37 (12), 5073–5087.

Karstanje, D., van der Wel, M., van Dijk, D., 2015. Common factors in commodity futures curves. Working paper.

Keynes, J., 1930. Treatise on money. London: Macmillan.

King, M., Sarno, L., Sojli, E., 2010. Timing exchange rates using order flow: The case of the loonie. Journal of Banking and Finance 34, 2917–2928.

Koopman, S., Durbin, J., 2000. Fast filtering and smoothing for multivariate state space models. Time Series Analysis 21, 281–296.

Koopman, S. J., Mallee, M. I. P., Van der Wel, M., 2010. Analyzing the term structure of interest rates using the dynamic Nelson-Siegel model with time-varying parameters. Journal of Business and Economic Statistics 28 (3), 329–343.

Kose, M. A., Otrok, C., Whiteman, C. H., 2003. International business cycles: World, region, and country-specific factors. The American Economic Review 93 (4), 1216–1239.

Ledoit, O., Wolf, M., 2008. Robust performance hypothesis testing with the sharpe ratio. Journal of Empirical Finance 15 (5), 850 – 859.

Lesmond, D., 2005. Liquidity of emerging markets. Journal of Financial Economics 77, 411–452.

Lesmond, D., Ogden, J. P., Trzcinka, C. A., 1999. A new estimate of transaction costs. Review of Financial Studies 12 (5), 1113–1141.

Liu, W., 2006. A liquidity-augmented capital asset pricing model. Journal of Financial Economics 82, 631–671.

Locke, P., Venkatesh, P., 1997. Futures market transaction costs. Journal of Futures markets 17 (2), 229–245.

Lutzenberger, F., 2014. The predictability of aggregate returns on commodity futures. Review of Financial Economics 23 (3), 120–130.

Marquering, W., Verbeek, M., 2004. The economic value of predicting stock index returns and volatility. Journal of Financial and Quantitative Analysis 39 (2), 407–429.

Marshall, B. R., Nguyen, N. H., Visaltanachoti, N., 2012. Commodity liquidity measurement and transaction costs. Review of Financial Studies 25 (2), 599–638.

Miffre, J., 2012. Comparing first, second and third generation commodity indices. Working paper.

Miffre, J., Rallis, G., 2007. Momentum strategies in commodity futures markets. Journal of Banking & Finance 31 (6), 1863–1886.

Milonas, N. T., 1991. Measuring seasonalities in commodity markets and the half-month effect. Journal of Futures Markets 11 (3), 331–345.
URL http://dx.doi.org/10.1002/fut.3990110307

Moskowitz, T., Ooi, Y., Pedersen, L., 2012. Time series momentum. Journal of Financial Economics 104 (2), 228–250.

Mouakhar, T., Roberge, M., 2010. The optimal approach to futures contract roll in commodity portfolios. The Journal of Alternative Investments 12 (3), 51–60.

Nelson, C., Siegel, A., 1987. Parsimonious modeling of yield curves. Journal of Business 60, 473–489.

Ohana, S., 2010. Modeling global and local dependence in a pair of commodity forward curves with an application to the US natural gas and heating oil markets. Energy Economics 32, 373–388.

O'Hara, M., 1995. Market Microstructure Theory. Blackwell, Cambridge, MA.

Pástor, Ł., Stambaugh, R. F., 2003. Liquidity risk and expected stock retums. The Journal of Political Economy 111 (3), 642–685.

Pesaran, M. H., Pick, A., 2011. Forecast combination across estimation windows. Journal of Business and Economic Statistics 29 (2), 307 – 318.

Pindyck, R., Rotemberg, J., 1990. The excess co-movement of commodity prices. Economic Journal 100, 1173–1189.

Pindyck, R. S., 2004. Volatility and commodity price dynamics. Journal of Futures Markets 24 (11), 1029–1047.

Pirrong, C., 2005. Momentum in futures markets. Working paper, EFA 2005 Moscow Meetings Paper.

Rime, D., Sarno, L., Sojli, E., 2010. Exchange rate forecasting, order flow and macroeconomic information. Journal of International Economics 80 (1), 72–88.

Roll, R., 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. Journal of Finance 39 (4), 1127–1139.

Sadka, R., 2006. Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk. Journal of Financial Economics 80 (2), 309 – 349.

Samuelson, P. A., 1965. Proof that properly anticipated prices fluctuate randomly 6 (2), 41–49.

Sargent, T. J., Sims, C. A., 1977. Business cycle modeling without pretending to have too much a priori economic theory. Working paper.

Schwartz, E. S., 1997. The stochastic behavior of commodity prices: Implications for valuation and hedging. Journal of Finance 52 (3), 923–973.

Schwartz, E. S., Smith, J. E., 2000. Short-term variations and long-term dynamics in commodity prices. Management Science 46 (7), 893–911.

Shen, Q., Szakmary, A. C., Sharma, S. C., 2007. An examination of momentum strategies in commodity futures markets. Journal of Futures Markets 27 (3), 227–256.

Shumway, T., 1997. The delisting bias in CRSP data. Journal of Finance 52 (1), 327–340.

Singleton, K. J., 2014. Investor flows and the 2008 boom/bust in oil prices. Management Science 60 (2), 300–318.

Sorensen, C., 2002. Modeling seasonality in agricultural commodity futures. Journal of Futures Markets 22 (5), 393–426.

Stock, J. H., Watson, M. W., 1989. New indexes of coincident and leading economic indicators. In: NBER Macroeconomics Annual 1989, Volume 4. MIT Press, pp. 351–409.

Stock, J. H., Watson, M. W., 2012. Generalized shrinkage methods for forecasting using many predictors. Journal of Business and Economic Statistics 30 (4), 481–493.

Stoll, H. R., Whaley, R. E., 2010. Commodity index investing and commodity futures prices. Journal of Applied Finance 20 (1), 7–46.

Szakmary, A. C., Shen, Q., Sharma, S. C., 2010. Trend-following trading strategies in commodity futures: A re-examination. Journal of Banking & Finance 34 (2), 409–426.

Szymanowska, M., de Roon, F., Nijman, T., van den Goorbergh, R., 2014. An anatomy of commodity futures risk premia. Journal of Finance 69 (1), 453–482.

Tang, K., Xiong, W., 2012. Index investment and financialization of commodities. Financial Analysts Journal 68 (6), 54–74.

Telser, L., 1958. Futures trading and the storage of cotton and wheat. Journal of Political Economy 66, 233–255.

Thornton, D. L., Valente, G., 2012. Out-of-sample predictions of bond excess returns and forward rates: An asset allocation perspective. Review of Financial Studies 25 (10), 3141 – 3168.

Vansteenkiste, I., 2009. How important are common factors in driving non-fuel commodity prices? A dynamic factor analysis. Working paper, European Central Bank.

Vasicek, O., 1977. An equilibrium characterization of the term structure. Journal of Financial Economics 5, 177–188.

Vayanos, D., 1998. Transaction costs and asset prices: A dynamic equilibrium model. Review of Financial Studies 11 (1), 1 – 58.

Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies 21 (4), 1455–1508.

West, J., 2012. Long-dated agricultural futures price estimates using the seasonal Nelson-Siegel model. International Journal of Business and Management 7 (3), 78–93.

West, K. D., Edison, H., Cho, D., 1993. A utility based comparison of some models for exchange rate volatility. Journal of International Economics 35 (1-2), 23–45.

Working, H., 1948. Theory of the inverse carrying charge in futures markets. Journal of Farm Economics 30 (1), 1–28.

Working, H., 1949. The theory of price storage. American Economic Review 39 (6), 1254–1262.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

556. C.L. BEHRENS, *Product differentiation in aviation passenger markets: The impact of demand heterogeneity on competition*
557. G. SMRKOLJ, *Dynamic Models of Research and Development*
558. S. PEER, *The economics of trip scheduling, travel time variability and traffic information*
559. V. SPINU, *Nonadditive Beliefs: From Measurement to Extensions*
560. S.P. KASTORYANO, *Essays in Applied Dynamic Microeconometrics*
561. M. VAN DUIJN, *Location, choice, cultural heritage and house prices*
562. T. SALIMANS, *Essays in Likelihood-Based Computational Econometrics*
563. P. SUN, *Tail Risk of Equidity Returns*
564. C.G.J. KARSTEN, *The Law and Finance of M&A Contracts*
565. C. OZGEN, *Impacts of Immigration and Cultural Diversity on Innovation and Economic Growth*
566. R.S. SCHOLTE, *The interplay between early-life conditions, major events and health later in life*
567. B.N.KRAMER, *Why don't they take a card? Essays on the demand for micro health insurance*
568. M. KILIÇ, *Fundamental Insights in Power Futures Prices*

569. A.G.B. DE VRIES, *Venture Capital: Relations with the Economy and Intellectual Property*

570. E.M.F. VAN DEN BROEK, *Keeping up Appearances*

571. K.T. MOORE, *A Tale of Risk: Essays on Financial Extremes*

572. F.T. ZOUTMAN, *A Symphony of Redistributive Instruments*

573. M.J. GERRITSE, *Policy Competition and the Spatial Economy*

574. A. OPSCHOOR, *Understanding Financial Market Volatility*

575. R.R. VAN LOON, *Tourism and the Economic Valuation of Cultural Heritage*

576. I.L. LYUBIMOV, *Essays on Political Economy and Economic Development*

577. A.A.F. GERRITSEN, *Essays in Optimal Government Policy*

578. M.L. SCHOLTUS, *The Impact of High-Frequency Trading on Financial Markets*

579. E. RAVIV, *Forecasting Financial and Macroeconomic Variables: Shrinkage, Dimension reduction, and Aggregation*

580. J. TICHEM, *Altruism, Conformism, and Incentives in the Workplace*

581. E.S. HENDRIKS, *Essays in Law and Economics*

582. X. SHEN, *Essays on Empirical Asset Pricing*

583. L.T. GATAREK, *Econometric Contributions to Financial Trading, Hedging and Risk Measurement*

584. X. LI, *Temporary Price Deviation, Limited Attention and Information Acquisition in the Stock Market*

585. Y. DAI, *Efficiency in Corporate Takeovers*

586. S.L. VAN DER STER, *Approximate feasibility in real-time scheduling: Speeding up in order to meet deadlines*

587. A. SELIM, *An Examination of Uncertainty from a Psychological and Economic Viewpoint*

588. B.Z. YUESHEN, *Frictions in Modern Financial Markets and the Implications for Market Quality*

589. D. VAN DOLDER, *Game Shows, Gambles, and Economic Behavior*

590. S.P. CEYHAN, *Essays on Bayesian Analysis of Time Varying Economic Patterns*

591. S. RENES, *Never the Single Measure*

592. D.L. IN 'T VELD, *Complex Systems in Financial Economics: Applications to Interbank and Stock Markets*

593. Y.YANG, *Laboratory Tests of Theories of Strategic Interaction*

594. M.P. WOJTOWICZ, *Pricing Credits Derivatives and Credit Securitization*

595. R.S. SAYAG, *Communication and Learning in Decision Making*

596. S.L. BLAUW, *Well-to-do or doing well? Empirical studies of wellbeing and development*

597. T.A. MAKAREWICZ, *Learning to Forecast: Genetic Algorithms and Experiments*

598. P. ROBALO, *Understanding Political Behavior: Essays in Experimental Political Economy*

599. R. ZOUTENBIER, *Work Motivation and Incentives in the Public Sector*

600. M.B.W. KOBUS, *Economic Studies on Public Facility use*

601. R.J.D. POTTER VAN LOON, *Modeling non-standard financial decision making*

602. G. MESTERS, *Essays on Nonlinear Panel Time Series Models*

603. S. GUBINS, *Information Technologies and Travel*

604. D. KOPÁNYI, *Bounded Rationality and Learning in Market Competition*

605. N. MARTYNOVA, *Incentives and Regulation in Banking*