FACULTEIT ECONOMIE
EN BEDRIJFSKUNDE

UNIVERSITEIT
GENT

Microeconomics of Financial Markets

Gert Elaut

2017

Dissertation submitted to the Faculty of Economics and Business Administration, Ghent University, in fulfillment of the requirements for the degree of Doctor in Economics

Microeconomics of Financial Markets

Gert Elaut

Doctoral Jury

Prof. dr. Patrick Van Kenhove	Dean-President, Ghent University
Prof. dr. Michael Frömmel	Promotor, Ghent University
Prof. dr. Dries Heyman	Ghent University
Prof. dr. Koen Inghelbrecht	Ghent University
dr. Alexander Mende	RPM Risk & Portfolio Management AB
dr. Kathryn Kaminski	MIT Sloan School of Management

To my parents and Gwen

Acknowledgements

This doctoral dissertation has been in the making for almost 6 years now, and even though doing research is largely an individual endeavor, I would not have made if it wasn't for the support of a number of people and institutions. Now that my PhD project is finally completed, I would like to take the opportunity to express my gratitude.

First of all, I would like to express my thanks to my PhD supervisor prof. dr. Michael Frömmel. I want to thank you for taking a leap of faith in hiring me. I would not be writing these words if you hadn't given me this unique opportunity. You allowed me to find my own way, to define my own research agenda, and to attend any conference our work got accepted to. Thank you for that.

I would also like to express my appreciation for the other members of my doctoral jury, dr. Kathryn Kaminski (MIT Sloan School of Management), dr. Alexander Mende (RPM Risk & Portfolio Management AB), prof. dr. Koen Inghelbrecht (Ghent University), and prof. dr. Dries Heyman (Ghent University) for taking a keen interest in my work. They scarified some of their precious time to read and comment on my writings in the weeks leading up to the pre-defense.

I would also like to thank the people of RPM Risk & Portfolio Management AB, and in particular Alexander Mende (again), John Sjödin, and Péter Erdős. Their invaluable input, both for the projects in which they were co-author as well as projects in which they were not actively involved, really helped me to take our work to the next level. They have also taught me a lot about the internal workings of the Managed Futures industry and the asset management industry in general.

In this context, I would also like to acknowledge the financial support from the European Union (EU) as part of the Marie Curie action in the European Commission's 7th FP, Industry-Academia Partnerships and Pathways Program (Project reference:324440) which allowed me, amongst other things, to visit RPM Risk & Portfolio Management AB for two research visits.

A special thank you goes out to my colleagues from the Department of Financial Economics for the extremely nice working environment. A lot of different elements (lunches at 12 a.m. sharp together, coffee breaks at 10 a.m. and 3 p.m., crazy Friday lunches, and department activities just to name a few) helped create a unique atmosphere that contributed positively to my mental well-being during the Ph.D. A special thanks goes to Nora, Frederick, Joris, Maria, Selien, and Martien who helped me find my way during the early days of the Ph.D. Turning to the colleagues that started working at the department at the same time or after me, I would like to thank my co-author Kevin, Jasmien, Thomas, Frederik, Elien, Nicolas D., Hannes, Wouter, Xing, Murat, Nicolas S., Satrajit, Bruno, and Marco. A special thanks also goes out to the head of the department Rudi, whose course on banking and finance in the summer of 2008 sparked my interest for finance. Finally, I would like to thank Nathalie and Sabine. Not just for their great administrative support, but also for being there for a chat.

I would also like to mention my fellow students from the Bachelor and Master in Economics and the Master in Banking and Finance (MBF). Especially the MBF colleagues ensured a good work hard, play hard balance.

Last but certainly not least, I would like to thank my parents, and my sister Gerlinde, who have been supporting me all my life and stood by my side at every stage of my personal and professional endeavors. My parents gave me all possible opportunities for self-development and always gave the highest priority to my education. A special thank you goes out to Gwen, whom I have been sharing my life with. Thank you for your love and support.

Gert Elaut

Ghent, June 9th 2017

Nederlandstalige samenvatting

Dit doctoraatsproefschrift is een verzameling van vijf essays die bijdragen leveren tot de literatuur rond alternatieve beleggingen. Hoewel elk van de hoofdstukken grotendeels op zichzelf staat, is er desondanks een duidelijke rode draad doorheen de verschillende hoofdstukken. Het is namelijk zo dat elk van de hoofdstukken verband houdt met specifieke beleggingsstrategieën en een fondsenindustrie ('Managed Futures' of 'Commodity Trading Advisors') die dit soort strategieën in de praktijk toepast.

In het eerste hoofdstuk onderzoeken we de kostenstructuur van Managed Futures fondsen en hefboomfondsen in het algemeen. We bekijken de typische kostenstructuur van hefboomfondsen, welke inhoudt dat de fondsbeheerder 2% beheerskosten per jaar aanrekent en een prestatievergoeding gelijk aan 20% van de gerealiseerde meerwaarde. In dit hoofdstuk onderzoeken we een andere dimensie van deze kostenstructuur. Er is namelijk ook een belangrijke tijdsdimensie verbonden met het aanrekenen van de prestatievergoeding, welke varieert over verschillende fondsen. Deze verborgen dimensie van de kostenstructuur, genaamd de 'crystallization frequency', heeft een economisch significante invloed op de kosten die beleggers betalen.

In hoofdstuk twee onderzoeken we de implicaties van lage maar persistente autocorrelatie in de maandelijkse rendementen van Managed Futures fondsen voor portefeuillebeheer. We argumenteren dat de geobserveerde autocorrelatie wellicht niet het resultaat is van illiquiditeit in de onderliggende posities, gezien de liquiditeit van de effecten die Managed Futures fondsen verhandelen. In de plaats daarvan hypothetiseren we dat deze autocorrelatie consistent is met een strategie die vaak kleine verliezen incasseert en occasioneel grote winsten genereert. We bevestigen deze hypothese empirisch en tonen aan dat de positieve autocorrelatie consistent is met het divergent risicogedrag van trendvolgende strategieën. We tonen verder ook aan dat Managed Futures fondsen die positieve autocorrelatie vertonen beter presteren dan fondsen met negatieve autocorrelatie. Het geobserveerd rendement kan wellicht niet verklaard worden door een concentratie in bepaalde strategieën, grootte en leeftijd van de fondsen, en vertekeningen in de dataset. Bovendien heeft positieve autocorrelatie geen negatieve impact op de diversificatievoordelen van Managed Futures fondsen.

Het derde hoofdstuk focust op een trendvolgende strategie in de context van high-frequency data. We onderzoeken met name de mogelijke oorzaken van een fenomeen dat bekend staat als 'intraday momentum', welke gedefinieerd wordt als een significant positief verband tussen het rendement in het eerste half uur van de handelsdag en rendement in het laatste half uur van de handelsdag. Met behulp van transactie-level data van de Moscow Interbank Currency Exchange (MICEX) voor het Russische Roebel-Amerikaanse Dollar over de periode 2005-2014 analyseren we de door de literatuur voorgestelde mogelijke oorzaken. Onze resultaten suggereren dat, voor de periode 2005-2014, intraday momentum in de Roebelmarket wellicht het gevolg is van risicoaversie onder marktmakers voor het aanhouden van posities buiten de handelsuren. Onze resultaten bevestigen verder ook eerdere bevindingen die suggereren dat expliciete handelsuren van belang zijn voor intraday momentum en dat het effect sterker is tijdens crisissen.

In het vierde hoofdstuk dragen we bij tot de literatuur rond de market timing. We onderzoeken, aan de hand van vertrouwelijke data voorzien door RPM Risk & Portfolio Management, of Managed Futures fondsen in staat zijn trends in financiële markten te anticiperen. We verbeteren de bestaande methodologie en gebruiken data van een hogere frequentie om de analyse uit te voeren, en vinden dat Managed Futures fondsen inderdaad een significant market timing talent hebben.

In hoofdstuk vijf proberen we de strategieën die Managed Futures fondsen gebruiken, te ontrafelen. Aan de hand van data met betrekking tot de financiële derivaten die deze fondsen verhandelen, implementeren we een trendvolgende strategie. In dit hoofdstuk proberen we dus de vertrouwelijke modellen gebruikt door deze fondsen zo nauw mogelijk te repliceren. Hierbij combineren we handelssignalen over een groot aantal tijdsbestekken om op die manier de sterkte van een trend te incorporeren in het beleggingsproces. We tonen aan dat de voorgestelde strategie de kenmerken van Managed Futures fondsen goed repliceert. De door ons voorgestelde strategie is bijgevolg een goede maatstaf voor het analyseren van kandidaat-fondsen.

Chapter 1

Crystallization – the Hidden Dimension of CTAs' Fee Structure[1]

Financial Analysts Journal
July/August 2015, Vol. 71, No. 4: 51–62.

1.1 Introduction

The impact of the two components of hedge funds' and Commodity Trading Advisors' (CTAs) fee structure, the incentive fee and the high-water mark clause, on hedge fund behavior has been discussed extensively in the academic literature. Especially their effect on fund managers' risk-taking behavior has received considerable attention.[2] However, the fee structure also has more direct consequences for investors, apart from changing the risk profile of the investment. Fees impact long-term wealth and investors are more and more starting to realize this, not in the least because of the current low yield environment. Consequently, hedge funds' fees are now subject to closer scrutiny and are negotiated more often than in the past.

[1]This chapter is based on joint work with John Sjödin (RPM Risk & Portfolio Management and Ghent University) and Michael Frömmel (Ghent University).

[2]Studies include Goetzmann, Ingersoll, and Ross (2003), Hodder and Jackwerth (2007), Kouwenberg and Ziemba (2007), Panageas and Westerfield (2009), and Agarwal, Daniel, and Naik (2009).

To illustrate the downward pressure on hedge funds' headline fee levels, we report in Table 1 the management fee and incentive fee of newly launched CTAs reporting to BarclayHedge. The Table illustrates that, while there has been no significant change in incentive fee levels, average management fee levels have been decreasing steadily over time.

A 2/20-fee structure, i.e. a management fee of 2% of assets under management combined with an incentive fee of 20% of gains, is and has been the standard cost for allocations in the hedge fund industry. It is generally supplemented with a high-water mark, such that investors only pay the incentive fee once any previous underperformance has been made up for.

However, headline fee levels are only one aspect of the fee structure that should be considered. Another element usually not taken into consideration when discussing hedge funds' fees, is the frequency at which a fund charges the incentive fee and updates its high-water mark. This feature is commonly referred to as the *crystallization frequency* or the *incentive fee payment schedule*.

The crystallization frequency differs from the accrual schedule, which is the schedule used to calculate and charge the fee to the fund's profit and loss account. Whereas the process of fee accrual does not impact investor returns, the same is not true for the fee crystallization. As the incentive fee crystallization frequency increases, the expected total fee load charged by the hedge fund manager increases as well.

To illustrate the above concepts, we provide a brief numerical example in Table 1.2. For simplicity, we consider a fee structure that consists of a 20% performance fee but no management fee.

This example shows how an identical gross performance leads to widely different performance fee loads when we vary the crystallization frequency. From the example the reader can easily infer the source of this difference in fee load; under quarterly crystallization, some of the fund's interim highs are allowed to materialize into performance fees. In the case of annual crystallization however, only the asset value at the end of the year matters.

In this article, we contribute to the understanding of hedge funds' fee structure in that we highlight and analyse the impact of the crystallization frequency on hedge funds' fee load. To the authors' best knowledge, no study has yet investigated this aspect to hedge funds' fee structure. This finding is compelling. The crystallization frequency forms the basis for the incentive fee calculation and the way hedge funds update their high-water mark. Consequently, it has a material effect on the fees investors

Table 1.1: Evolution in CTA Headline Fee Levels

	Number of Funds	Management Fee	Bootstrapped 95% CI	Incentive Fee	Bootstrapped 95% CI
Prior to 1994	387	2.25%	[2.14%;2.36%]	20.38%	[20.09%;20.66%]
1994-1998	295	1.97%	[1.88%;2.06%]	20.63%	[20.29%;20.97%]
1999-2004	394	1.71%	[1.65%;1.78%]	20.51%	[20.24%;20.81%]
2005-2008	377	1.67%	[1.6%;1.73%]	20.71%	[20.3%;21.16%]
2009-2012	163	1.62%	[1.51%;1.72%]	20.64%	[19.9%;21.43%]
1994-2012	1616	1.87%	[1.83%;1.91%]	20.56%	[20.39%;20.74%]

This table reports summary statistics on the evolution of headline fee levels. In particular, we report the number of newly launched funds and the average incentive- and management fee for CTAs in BarclayHedge for the different sub-periods.

Table 1.2: Illustration Effect of Crystallization

		Annual Crystallization				Quarterly Crystallization			
Time	Gross Return	HWM	Incentive Fee Accrued	Incentive Fee Paid	NAV	HWM	Incentive Fee Accrued	Incentive Fee Paid	NAV
Jan	1.3%	100	0.26		101.30	100	0.26		101.30
Feb	0.3%	100	0.32		101.60	100	0.32		101.60
Mar	3.2%	100	0.97		104.86	100	0.97	0.97	103.88
Apr	3.6%	100	1.73		108.63	103.88	0.75		107.62
May	-0.9%	100	1.53		107.65	103.88	0.55		106.66
Jun	3.0%	100	2.18		110.88	103.88	1.19	1.19	108.66
Jul	-2.2%	100	1.69		108.44	108.66	0.00		106.27
Aug	-1.5%	100	1.36		106.82	108.66	0.00		104.68
Sep	0.0%	100	1.36		106.82	108.66	0.00	0.00	104.68
Oct	-0.9%	100	1.17		105.85	108.66	0.00		103.73
Nov	-2.3%	100	0.68		103.42	108.66	0.00		101.35
Dec	1.8%	100	1.06	1.06	104.23	108.66	0.00	0.00	103.17

This table reports the fees paid by an investor under annual and quarterly crystallization, respectively. The initial HWM and NAV equal 100. The fee structure in this example equals 0/20%, i.e. no management fee and a performance fee of 20% of realized gains.

pay and could also influence hedge funds' risk-taking behavior.

Our findings have several implications, both for researchers and practioners. First, we show that the choice of the crystallization frequency has both a statistically and economically significant impact on fees paid by investors. In the case of CTAs, and assuming a 2/20-fee structure, shifting from annual to quarterly crystallization leads to a 49 basis points increase in the annual fee load (as a percentage of assets under management). In addition, an incentive fee of 15% combined with monthly crystallization leads to the same total fee load as an incentive fee of 20% under annual crystallization. Both results imply that the effect of the crystallization frequency is important for allocators evaluating and comparing different fund investments. We stress that, while we focus on just one hedge fund category, CTAs, the crystallization frequency is an important consideration in any investment vehicle whose fee structure depends on a high-water mark provision. Moreover, in an environment where especially hedge funds' management fee levels are under pressure, the relative importance of the incentive fee and, thus, crystallization in the total fee load increases.

Second, our study also has implications for academic literature that estimates hedge funds' gross returns and fee loads as well as research on hedge funds' risk-taking behavior. To construct gross returns, previous studies in most cases assume that incentive fees are paid at year-end (e.g. Brooks, Clare, and Motson (2007), French (2008) and Agarwal, Daniel, and Naik (2009)), although some authors assume quarterly payment (see Bollen and Whaley (2009) and Jorion and Schwarz (2014)). Certain authors also calculate hedge funds' historical fee load in their analysis. French (2008) estimates that the typical investor in U.S. equity-related hedge funds has paid an annual combined fee or total expense ratio of 3.69% p.a. over the period 2000-2007. Brooks, Clare, and Motson (2007) find that between 1994 and 2006 hedge fund fees averaged 5.15% annually. Ibbotson, Chen, and Zhu (2011) suggest a lower estimate of 3.43% p.a. for the period 1995 to 2009. Similarly, Feng, Getmansky, and Kapadia (2011) report total fees over the period 1994-2010 to be on average 3.36% of gross asset value. However, these studies do not consider the impact of the crystallization frequency on these figures. With regard to hedge funds' risk-taking behavior, our analysis has implications for the time frame over which previous results on hedge funds' risk-taking behavior might apply. If fund managers update their high-water mark more than once a year, their trading horizon is shortened accordingly.

Finally, crystallization frequencies of hedge funds have not been documented previously. To shed light on crystallization practices, we perform a survey among the constituents of the Newedge CTA Index as well as an analysis of the fee notes of CTAs in the Tremont Advisory Shareholder Services (TASS) database. We find that, at least in the case of CTAs, high-water marks are most often updated quarterly, rather than annually. These findings for the CTA hedge fund category contrast the view commonly held in the academic literature that the high-water marks in hedge funds are commonly set at the end of the year.

For completeness, we focus on the impact of the crystallization frequency of the incentive fee, and we do not go into the payment frequency of the management fee. We do this mainly because the payment of the management fee does not depend on a fund's high-water mark.[3]

1.2 Data

We analyse the impact of the crystallization frequency on fees paid by investors by using monthly net-of-fee returns of live and dead funds labelled CTA in the BarclayHedge Database. We use a sample that covers the period January 1994 to December 2012 to mitigate a potential survivorship bias, since most databases only started collecting information on defunct programs from 1994 onwards.[4] As BarclayHedge does not report a first reporting date, we cannot eliminate the backfill bias entirely. We therefore opt for an alternative approach and remove the first twelve observations of a fund's return history, following Teo (2009).[5,6]

We further require at least twelve return observations for a fund to be included, and only include funds whose monthly returns are denominated in USD or EUR.[7] The EUR-denominated returns are converted to USD-

[3]In addition, the vast majority of the funds charge the management fee monthly. For the Tremont Advisory Shareholder Services (TASS) database, we find that 78% of the CTAs in the database charge the management fee on a monthly basis. 13% charges the management fee quarterly and 8% charges the management fee annually.

[4]Gross returns are first calculated using the funds' entire return history, after which the pre-1994 period is dropped.

[5]We first calculate gross returns (see Section 1.4.1) using the fund's entire track record, and afterwards drop the first twelve observations of the fund's net-of-fee and gross returns.

[6]By keeping track of the amount of months that are backfilled when a fund is first included to BarclayHedge database, we tracked backfill bias for the period 2005-2010. For that sample period, the median (average) backfill bias was twelve (fourteen) months.

[7]Programs denominated in currencies other than USD and EUR are in most instances duplicate share classes of larger programs and would therefore be dropped in

denominated returns, using the end-of-month spot USD/EUR exchange rate. As the analysis also requires information on the funds' management fee and incentive fee, we remove cases where at least one of the two variables is unreported.[8]

We then filter the resulting sample of funds by looking at their self-declared strategy description and remove funds whose description is not consistent with the definition of CTAs. In the process, we also determine whether the program under consideration is the fund's flagship program and discard duplicates. To ensure that our results apply to funds that can be considered part of the investable universe for most CTA investors, we remove funds whose net-of-fee returns exhibit unusually low- or high levels of variation. To this end, we discard funds when the standard deviation of the observed net-of-fee returns is lower than 2% or exceeds 60% p.a. After applying these restrictions, our sample consists of 1,616 unique CTA programs. Table 1.3 reports summary statistics for the final set of funds.

Table 1.3: Summary Statistics CTAs

	Mean	Min	P25	P50	P75	Max
Monthly net-of-fee return	0.57%	-6.47%	0.06%	0.50%	0.99%	9.52%
Monthly standard deviation	5.08%	0.61%	2.75%	4.27%	6.59%	17.17%
Age (years)	5.4	1	2.1	3.8	7	19
Management fee	1.87%	0%	2%	2%	2%	5%
Incentive fee	20.56%	5%	20%	20%	20%	50%

This table reports summary statistics for the sample of 1616 CTAs from the Barclay-Hedge database.

In this paper, we focus on one hedge fund category and CTAs in particular because industry standards on crystallization for different hedge fund categories might differ. It is possible that the crystallization frequency of hedge funds is to some extent related to differences in the ability of funds to value their underlying positions. Unlike some other hedge fund categories, CTAs trade almost exclusively highly liquid instruments and, thus, do not have any practical limitations regarding the calculation of NAVs. As such, CTAs provide a fruitful ground for analysing the impact of crystallization.

any case.

[8]Additionally, we also exclude cases where both types of fee are zero or and cases where the fee levels are deemed unreasonable low or high (management fee in excess of 5% p.a., incentive fees below 5% or above 50% p.a.).

1.3 Crystallization and Industry Practices

Since public hedge fund databases do not keep track of funds' incentive fee crystallization frequency[9], we perform a survey among the constituents of the Newedge CTA index (as of May 2013). The Newedge CTA index is designed to track the largest CTAs and aims to be representative of the Managed Futures space. The index is comprised of the 20 largest managers (based on AUM) who are open to new investment and that report performance on a daily basis to Newedge. Where possible, we complete the results of the survey with information available on the website of the U.S. Securities and Exchange Commission (SEC).[10]

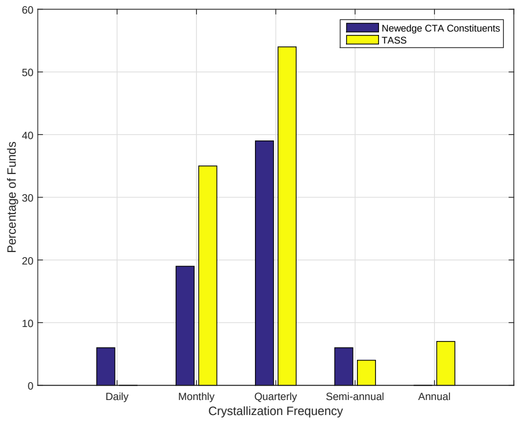
The results of the survey are reported in Figure 1.1. The bar chart indicates that, in the case of CTAs, the most commonly used crystallization frequency is quarterly. In those instances where the crystallization frequency is not quarterly, we find that the frequency generally tends to be higher, rather than lower. In unreported results, we weigh the crystallization frequency by the assets under management (AUM) of every manager. While quarterly crystallization remains the most commonly applied crystallization frequency (55% of AUM), monthly crystallization increases in importance as it applies to 28.3% of AUM covered by the survey. Finally, to gauge the scope of our survey vis-à-vis total AUM by the CTA industry, the results of our survey cover 57% of assets managed in the CTA space that report to BarclayHedge.

As mentioned above, public databases do not keep track of the crystallization frequency in a systematic way. However, the fee notes in the Tremont Advisory Shareholder Services (TASS) database in a number of cases do provide a sufficient amount of information to pinpoint the crystallization frequency. Therefore, and in addition to the above survey, we also examine the fee notes of defunct and live CTAs reported in the TASS database. The results are also reported in Figure 1.1. Comparing these results with those of our own survey suggests that the sample of funds from TASS is characterised by higher crystallization frequencies. These differences could be due to survivorship bias as well as differences in fund size. Nevertheless, the results for the TASS sample corroborate our ear-

[9] TASS's questionnaire only inquires about the management fee's payment frequency; the other widely used databases' questionnaires and manuals (Hedge Fund Research (HFR), CISDM, and BarclayHedge) indicate that the databases do not keep track of the fee payment frequencies.

[10] In particular, we make use of the SEC's Investment Adviser Public Disclosure (IAPD) and the Electronic Data-Gathering, Analysis, and Retrieval (EDGAR) database.

Figure 1.1: Distribution of the Crystallization Frequencies of the Incentive Fee



lier finding that quarterly is the most common crystallization frequency. When funds use a crystallization frequency other than quarterly crystallization, the frequency tends to be higher rather than lower.

For completeness, we also look at the relationship between the reported fee levels and the crystallization frequency of the funds. It could be that funds with lower crystallization frequencies have higher incentive fee levels, such that the total fee load is comparable. To verify that this is not the case, we group the sample of funds in TASS based on their reported crystallization frequency and analyse the average incentive and management fee of the different groups. The results, reported in Table 1.4, indicate that funds with a higher crystallization frequency tend to have higher headline incentive fee levels. For example, the average incentive fee level for funds with monthly crystallization (22.38%) is significantly higher than that of funds that employ a quarterly crystallization frequency (21.05%), with a p-value of 0.0775. In addition, we also find that the headline management fee level tends to increase as the crystallization frequency increases. These results suggest that funds that apply higher a crystallization frequency on average also charge higher headline fee levels.

Table 1.4: Relationship between Crystallization Frequency and Fee Levels

	Incentive Fee	Bootstrapped 95% CI	Management Fee	Bootstrapped 95% CI
Monthly	22.38%	[20.72%;24.23%]	1.63%	[1.36%;1.91%]
Quarterly	21.05%	[20.35%;21.8%]	1.64%	[1.48%;1.79%]
Semi-annual	20.00%	[20%;20%]	1.93%	[1.79%;2%]
Annual	19.62%	[17.69%;21.15%]	1.47%	[1.17%;1.81%]

This table reports the average incentive fee level and management fee level under different crystallization frequencies for sample of CTAs in TASS.

1.4 Incentive Fee Crystallization and Fee Load

1.4.1 Construction of Gross Returns

As mentioned in the introduction, analysing the impact of the crystallization frequency on hedge funds' fee load requires calculating hedge funds' gross returns and charging fees to investors under various crystallization frequencies. To this end, we develop an algorithm that achieves this objective. We provide a thorough description of the algorithm in the Appendix.

To calculate gross returns for the sample of CTAs, we assume that CTAs apply quarterly crystallization to charge incentive fees. Our survey results and the results from TASS's fee notes suggest that this is the most commonly used crystallization frequency. In addition, when CTAs apply another crystallization frequency, they generally tend to use higher crystallization frequencies. As such, the assumption of quarterly fee crystallization should lead to fairly conservative estimates of the funds' gross returns.

In Table 1.5 we compare the observed net-of-fee CTA returns with the obtained gross CTA returns. Funds appear to earn significantly higher risk-adjusted returns – measured by the annualized Sharpe ratio – based on gross returns, as compared to net-of-fee returns. Also, both skewness and kurtosis are significantly higher for the gross returns. Consequently, we find a higher proportion of cases in which the Jarque-Bera test for normality rejects the null hypothesis of normality. Finally, we find that both net-of-fee returns and gross returns of CTAs exhibit negative first order serial correlation.

Table 1.5: Comparison of Net-of-fee Returns and Gross Returns

	Net-of-fee Returns	Gross Returns	p-value
Average return	0.57%	0.77%	0
Standard deviation of monthly returns	5.08%	4.68%	0
Annualized Sharpe Ratio	0.48	0.69	0
Skewness	0.31	0.45	0
Kurtosis	4.82	5.13	0.013
First order serial autocorrelation	-0.011	-0.004	0.138
JB-Statistic (Percentage of rejections)	47.22%	52.23%	

This table compares net-of-fee returns with the estimated gross returns based on the algorithm described above for the set of 1616 CTAs.
The reported p-values test the difference in means using the empirical t-distribution (bootstrap).

1.4.2 Analysis of the Historical Effect

As an introduction to our main analysis, we first estimate the crystallization frequency's potential historical effect on investor wealth. This way, we can get a feel of the economic significance of the effect of crystallization. Using the data set of gross returns obtained in Section 1.4.1, we re-apply the fund's reported headline fee levels under different crystallization frequencies. This way we obtain net-of-fee returns under different crystallization frequencies as well as the corresponding fee load.

In Table 1.6 we report the average gross return, average net-of-fee return, and the average fee load under the different fee crystallization schemes. The reported average net-of-fee returns are all statistically different from each other at the 1% level of significance (p-values unreported for conciseness). Furthermore, the results suggest that investors whose investment is subject to quarterly (monthly) crystallization, will earn net-of-fee returns which are on average 25 (42) basis points per year lower than in the case of annual crystallization. To put these figures into perspective, an annual difference of 42 basis points over a 10-year period will compound to a difference of 9.32% in the expected capital gain. For a MUSD 1 initial investment, this difference equals USD 63,303.

Even more important than these absolute numbers, is the impact on the risk-adjusted performance. Our results suggest that when investors move from annual to monthly crystallization, the Sharpe ratio deteriorates from 0.4 to 0.34, a 15.65% decrease.

We also observe from Table 1.6 that management fees are slightly lower than 2% p.a., despite the positive drift in CTAs their returns. This

Table 1.6: Summary Statistics Historical Fee-loads

	Average	Standard Deviation	Sharpe Ratio		
Gross Return	8.65%	16.22%	0.61		
	Net-of-fee Return	Standard Deviation	Sharpe Ratio	Management Fee	Incentive Fee
Monthly	4.90%	16.75%	0.34	1.93%	2.41%
Quarterly	5.07%	16.33%	0.37	1.93%	2.26%
Semi-annual	5.20%	16.05%	0.38	1.93%	2.16%
Annual	5.32%	15.75%	0.40	1.94%	2.14%

This table reports the average annual gross return, average standard deviation and average Sharpe ratio for the set of 1616 CTAs. The second part of the table reports the corresponding statistics for the net-of-fee returns, as well as the average management fee and incentive fee.

is consistent with our finding that management fees, at least for newly launched funds, tend to be below 2% p.a. on average (see Table 1).

1.4.3 Block Bootstrap Analysis

To study the effect of the crystallization frequency on the level of fees investors pay, we analyse the effect of crystallization by applying a block bootstrap. In particular, we randomly sample gross return histories and calculate the fee load under different crystallization regimes. The advantage of this approach is that we do not have to make any distribution assumptions with regard to the return generating process. A block bootstrap allows us to account for higher moments in monthly returns (e.g. CTAs' returns exhibit positive skewness) and to preserve any autocorrelation present in the gross return data. These properties of the return generating process can have a material impact on the results of the analysis and investors' total fee load.

In performing the block bootstrap, we consider all the potential 12/36/60-month samples in the data set of gross returns and pick 10,000 12-months, 36-month and 60-month samples. To avoid a potential look-ahead bias, we allow the sampling procedure to select incomplete samples occurring at the end of a fund's track record. In those cases where a fund terminates before the end of the sample period, we assume that investors redeem.[11]

[11]While most of these occurrences will correspond to fund terminations due to bad performance, we nevertheless treat the fund's exit as full redemption. If there is a

Table 1.7: Impact of Crystallization on Fee Load

Crystallization Frequency		Incentive Fee	Management Fee	Total Fee Load
1-year horizon	Monthly	2.76%***	2.07%**	4.84%***
	Quarterly	2.42%***	2.07%	4.50%***
	Semi-annual	2.19%***	2.08%	4.27%***
	Annual	1.93%	2.08%	4.01%
3-year horizon	Monthly	2.06%***	2.06%	4.13%***
	Quarterly	1.86%***	2.06%	3.93%***
	Semi-annual	1.73%***	2.06%	3.79%***
	Annual	1.61%	2.06%	3.67%
5-year horizon	Monthly	1.84%***	2.05%	3.89%***
	Quarterly	1.67%***	2.05%	3.72%***
	Semi-annual	1.55%***	2.05%	3.61%***
	Annual	1.44%	2.05%	3.50%

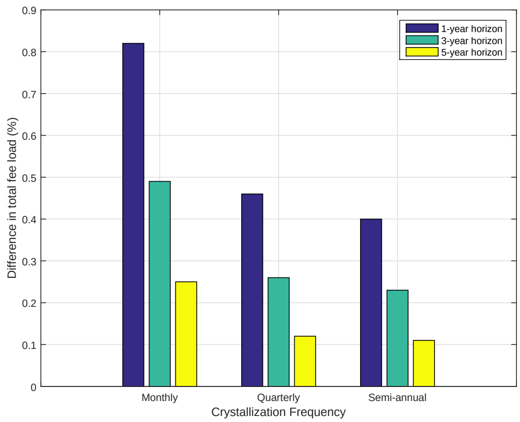
This table reports the average incentive fee, average management fee, and average total fee from performing a block bootstrap where 12, 36, or 60 month blocks of gross returns are drawn from the obtained sample of CTAs. Fee load equals the average annual fee load over the investment horizon, as a percentage of initial NAV/NAV at the end of the previous year.
Asterisks report statistically significance of the difference between of the obtained fee levels and the benchmark category (annual crystallization) at the 10% (*), 5% (**) and 1% (***) level of significance. Significance tests based on the empirical t-distribution (bootstrap).

We also assume that every draw starts the beginning of a calendar year (i.e. from January onwards). Having selected a random sample path of gross returns, we apply a standard 2/20-fee structure under different crystallization frequencies. This framework allows us to examine the impact of the crystallization frequency on investors' total fee load.

Table 1.7 reports the results for one-year, three-year, and five-year investment horizons. We consider periods of up to five years as this corresponds to the average age of the CTAs in the sample (see Table 1.3). As such, our analysis covers the relevant horizon over which the effect of crystallization applies for the majority of hedge fund investors. To gauge the significance of the results, we indicate whether the obtained fee level differs significantly from the fee load under annual crystallization. We

positive accrued interest fee at the time of the last observation, it will be charged to the investor's account.

Figure 1.2: Comparing the Total Fee Load with Annual Crystallization



set annual crystallization as the benchmark since most previous research made the assumption that the incentive fee is paid at the end of the year.

Our results illustrate that a higher crystallization frequency always leads to a higher average fee load.[12] Management fees are slightly higher than 2% and increasing in time due to the positive drift in the CTAs' returns. We find significantly higher fee loads as the crystallization frequency increases. The effect is also economically significant. For the one-year investment horizon, the total fee load is 49 (82) basis points p.a. higher in the case of quarterly (monthly) crystallization when compared to annual crystallization. This suggests that, under a 2/20-fee structure, the fee load is expected to be 12.2% (20.5%) higher if a manager charges the incentive fee quarterly (monthly), rather than annually. If the investment horizon is extended to five years, the difference decreases 23 (40) basis points p.a., a difference of 6.5% (11.4%). For ease of comparison and Figure 1.2 provides a graphical representation of the difference in fee load, with annual crystallization serving as the baseline.

[12]An alternative way to illustrate this finding, is by using option pricing. Indeed, the performance fee earned by the manager over any subperiod is a fraction (20%) of the value of a European call option with a strike price equal to the investor's HWM. Using Monte-Carlo simulation, it is easy to show that an exotic option, consisting of a sequence of European call options with path-dependent strike prices equal to the relevant HWM, is more valuable than a single European call option over the same period.

In addition to the increase in fee load as we increase the crystallization frequency, several other observations are evident from the results in Table 1.7. First, increasing the investment horizon dampens the impact of a higher crystallization frequency on fee load. We can explain this finding by the fact that the fee loads reported for the three- and five-year investment horizons are an average across the individual years. In years where a fund is not able to charge incentive fees, the total fee is the same under different crystallization frequencies. Despite this downward drag on the total fee load, caused by years in which only a management fee is paid, the difference in fee load for the different crystallization frequencies remains significant.

Second, for the one-year investment horizon, the management fee in the case of monthly crystallization is significantly lower than that under annual crystallization. This illustrates the fact that a higher crystallization frequency lowers the NAV on which funds can charge the management fee, since an incentive fee payment lower the investor's NAV. However, the effect is small in economic terms and more than offset by the higher fee load that results from the higher incentive fees paid.

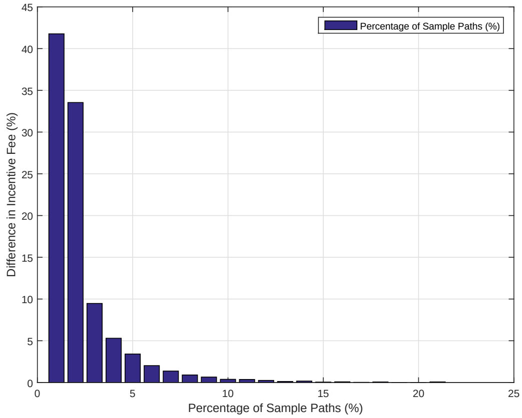
Next, we have a look at the distribution of the difference in fee loads. From the above analysis, we collect the set of differences in incentive fee under annual and quarterly crystallization. The results, reported in Figure 1.3, illustrate how the distribution of differences is highly skewed to the right.[13] The Figure also shows that in approximately 41.77% of the cases, the two crystallization frequencies do not show any difference in fee load. This is the case whenever (a) a fund does not get over its initial high-water mark, (b) when new highs are reached but not crystallized and (c) when the fund sets new high-water marks at every crystallization date.

In the first two instances, investors only pay the management fee, which is the same for both crystallization frequencies. Of course, investors invest with a positive view on the investment's future performance. An unintended consequence of a higher crystallization frequency is therefore that the investors will pay more (i.e. there will be a positive difference in the fee load) at times when investors are generally less satisfied with the fund's performance.

To see this, consider the following case. When a fund manager, during a particular year, performs very well and continuously sets new highs

[13]This particular distribution is also the reason is why all tests of statistical significance are done using an empirical t-distribution (bootstrap).

Figure 1.3: Distribution of Difference in Incentive Fee Load



until the end of the calendar year, it does not matter what crystallization frequency is applied. However, in cases where the fund's NAV at year-end drops below a high-water mark set during the year –the difference in fee load under different crystallization frequencies will be positive. In those cases, investors will be paying higher fees while at the same time the fund's newly crystallized high-water mark will actually be above the NAV at the end of the year (i.e. a drop in NAV). This makes it clear that a higher crystallization frequency will tend to decrease the fund manager's investment horizon and lower the incentive to perform subsequent to the crystallization.

When we condition on those bootstrapped cases where an incentive fee is actually payable, the difference in incentive fee load is 78 basis points higher under quarterly crystallization, as compared to annual crystallization. Comparing this result to the unconditional average, a 49 basis points difference, suggests that in those cases that investors actually pay an incentive fee, the fee load will be higher than our main results would suggest.

Table 1.8: Trade-off between Crystallization Frequency and Incentive Fee

	Incentive Fee (%)					
Crystallization Frequency	5	10	15	20	25	30
Monthly	2.57%	3.07%	3.60%	4.08%	4.61%	5.24%
Quarterly	2.53%	2.97%	3.46%	3.88%	4.36%	4.94%
Semi-annual	2.50%	2.91%	3.36%	3.75%	4.20%	4.73%
Annual	2.46%	2.84%	3.26%	3.62%	4.03%	4.53%

This table reports the total fee load under different combinations of the both negotiable factors, the incentive fee level and the crystallization frequency. The management fee is paid monthly and fixed at 2% p.a. The fee load is estimated by drawing random three-year sample paths from the gross CTA return data and calculating the fee load, varying the crystallization frequency and the level of the incentive fee.

1.4.4 Trade-off between Incentive Fee and Payment Frequency

So far, we have assumed a standard 2/20-fee structure to analyse the impact of different payment frequencies. The analysis has shown that, when investors want to compare the (expected) fee load between different funds, such a comparison will be inaccurate if funds differ in terms of the incentive fee payment frequency. In this subsection, we quantify the trade-off that exists between the incentive fee and the crystallization frequency, keeping fixed the level and payment frequency of the management fee. This trade-off might be relevant if the crystallization frequency and incentive fee level are considered negotiable factors.

To ensure that our obtained estimates of the fee load are close to what an investor can expect in reality, the figures are also based on the block bootstrap outlined above. In particular, we calculate the fee load for 10,000 randomly drawn three-year sample paths of gross returns and vary the crystallization frequency and the incentive fee level.

Table 1.8 reports the size of the effect for different combinations of both negotiable factors. Unlike what incentive fee headline levels would suggest, the table illustrates that changes in the crystallization frequency lead to considerable differences in total fee load. For example, the results suggest that a 15% incentive fee with monthly crystallization leads to a similar total fee load as a 20% incentive fee with annual crystallization (not significantly different).

1.5 Robustness Checks

We now perform a number of robustness checks with regard to the level of the effect. Relaxing or imposing additional restrictions on the dataset used in the analysis will not change our finding that higher crystallization frequencies increase investors' fee load. However, it might have an influence on level of the fee loads and the economic significance of the effect of crystallization.

1.5.1 Impact of Backfill Bias

In our baseline analysis we account for backfill bias by discarding the first twelve observations of a fund's track record. Here we investigate the importance of this assumption for our baseline results.

To this end, we perform the following analysis. We redo the bootstrap analysis used in section 1.4.3 a 100 times, both for the baseline gross return data set and the newly obtained gross return data that does not correct for backfill bias. Then, we test whether the results in both cases differ significantly. Panel A of Table 1.9 reports the result. In line with our expectations, we find that a potential backfill bias tends to upward bias the obtained incentive fee loads. Nevertheless, the size of the difference in fee loads remains similar in both instances, both in terms of magnitude and statistical significance.

1.5.2 Impact of Fund Size

Another possible concern, raised by Kosowski, Naik, and Teo (2007), is that funds with assets under management below MUSD 20 might be too small for many institutional investors. To ensure that the magnitude of fee load differences is representative and do not deviate too much from the fee load institutional investors can expect, we perform the following robustness check.

Similar to the previous robustness check, we redo the bootstrap analysis a 100 times, but impose an additional restriction when selecting a sample path. In particular, we only select a sample path if –at the start– the corresponding fund's assets under management are above MUSD 20. To avoid look-ahead bias, the fund's size is allowed to drop below MUSD 20 in subsequent months. Results are reported in panel B of Table 1.9. Consistent with the finding that small funds tend to outperform more

Table 1.9: Results Robustness Checks

Robustness check		Baseline Result	Result under Robustness Check	p-value
Backfill Bias	Monthly	4.11%	4.38%	0
	Quarterly	3.91%	4.17%	0
	Semi-Annual	3.78%	4.03%	0
	Annual	3.66%	3.89%	0
Fund Size	Monthly		3.65%	0
	Quarterly		3.49%	0
	Semi-Annual		3.37%	0
	Annual		3.26%	0
Risk-taking Behavior	Monthly		4.11%	0.48
	Quarterly		3.92%	0.07
	Semi-Annual		3.79%	0.04
	Annual		3.71%	0

This table reports the total fee load for a three-year investment horizon for the baseline case, and a set of three robustness checks.
The reported p-values test the difference in means using the empirical t-distribution (bootstrap).

mature funds, we find that the fee load is lower when we omit smaller funds.

1.5.3 Impact of Risk-taking Behavior

To perform the bootstrap in the baseline case, we assume that every sample path drawn from the gross return dataset starts in January. However, Nanda and Aragon (2012) show that hedge funds take part in tournament behavior. Hedge funds tend to increase their risk-profile in the second half of the year when they are underperforming, relative to their peers. As such, the funds' risk-profile could differ throughout the calendar-year, and thus have an impact on our reported fee loads. To check whether this is the case, we redo the bootstrap and select sample paths that correspond to actual calendar-years.

The results are reported in panel C of Table 1.9. The p-values in Panel C indicate that in most cases, the total fee load is somewhat higher if we use actual calendar-years. We interpret this finding as being in line with the results by Aragon and Nanda (2012) on risk-taking behavior among hedge funds. Our results indicate that, taking into account intra-year patterns in the funds' returns, we find higher total fee loads. This result therefore suggests that funds actively change their exposure to safeguard accrued incentive fees, causing our results to exhibit slightly higher fee loads if we take these intra-year patterns into account.

1.6 Conclusion

The fee load of investors does not depend on the headline fee levels alone. Other aspects of the fee structure should also be considered when analysing fee structures that include incentive fees and a high-water mark provision. One such factor is the frequency with which hedge funds update their high-water mark.

To the best of our knowledge we are the first to document the impact of the crystallization frequency on hedge funds' fee loads. Using a bootstrap based on a comprehensive data set of CTAs, our main finding is that, under a 2/20-fee structure, quarterly crystallization leads to a fee load which is on average 49 basis points p.a. higher than under annual crystallization. This difference is economically large and should be a relevant consideration when discussing the fee structure. Our results are relevant for allocators who want to assess the fee load of fee schemes which differ in terms of crystallization frequency. Moreover, we find that different headline fee levels can lead to similar total fee loads, once the crystallization frequency is taken into consideration.

A failure to take into account the frequency with which the high-water mark is updated leads to erroneous estimates of funds' gross returns. In particular, assuming an annual payment of the incentive fee when the industry standard of a number of hedge fund categories is akin to quarterly crystallization, will lead to the underestimation of the gross returns of those hedge fund categories. As such, while annual crystallization might be common among some hedge fund categories, we document that quarterly crystallization is the most common crystallization frequency among CTAs.

Our analysis of the crystallization frequency suggests several avenues for future research. First, we did not go into the implications of the payment frequency on the risk-taking behavior of hedge funds and CTAs. Changes in the crystallization frequency alter the horizon over which the implications of the high-water mark on risk-taking behavior should be evaluated. As such, it can be expected that a higher crystallization frequency leads to a shorter trading horizon, and thus might conflict with a fund's stated strategy horizon. Second, we only cover one hedge fund category. As such, there might be considerable differences in the crystallization frequencies applied by different hedge fund categories. These differences might be related to hedge fund characteristics such as the liquidity of the strategy.

Bibliography

Agarwal, Vikas, Naveen D. Daniel, and Narayan Y. Naik, 2009, Role of managerial incentives and discretion in hedge fund performance, *The Journal of Finance* 64, 2221–2256.

Aragon, George O, and Vikram Nanda, 2012, Tournament behavior in hedge funds: high-water marks, fund liquidation, and managerial stake, *Review of Financial Studies* 25, 937–974.

Bollen, Nicolas P.B., and Robert E. Whaley, 2009, Hedge fund risk dynamics: Implications for performance appraisal, *The Journal of Finance* 64, 985–1035.

Brooks, Chris, Andrew Clare, and Nicholas Motson, 2007, The gross truth about hedge fund performance and risk: the impact of incentive fees, *Available at SSRN* .

Feng, Shuang, Mila Getmansky, and Nikunj Kapadia, 2011, Flows: The 'invisible hands' on hedge fund management, in *Midwest Finance Association 2012 Annual Meetings Paper*.

French, Kenneth R., 2008, Presidential address: The cost of active investing, *The Journal of Finance* 63, 1537–1573.

Goetzmann, William N., Jonathan E. Ingersoll, and Stephen A. Ross, 2003, High-water marks and hedge fund management contracts, *The Journal of Finance* 58, 1685–1718.

Hodder, James E., and Jens Carsten Jackwerth, 2007, Incentive contracts and hedge fund management, *Journal of Financial and Quantitative Analysis* 42, 811–826.

Ibbotson, Roger G, Peng Chen, and Kevin X Zhu, 2011, The abcs of hedge funds: alphas, betas, and costs, *Financial Analysts Journal* 67, 15–25.

Jorion, Philippe, and Christopher Schwarz, 2014, Are hedge fund managers systematically misreporting? or not?, *Journal of Financial Economics* 111, 311–327.

Kosowski, Robert, Narayan Y. Naik, and Melvyn Teo, 2007, Do hedge funds deliver alpha? a bayesian and bootstrap analysis, *Journal of Financial Economics* 84, 229–264.

Kouwenberg, Roy, and William T. Ziemba, 2007, Incentives and risk taking in hedge funds, *Journal of Banking and Finance* 31, 3291–3310.

Panageas, Stavros, and Mark M. Westerfield, 2009, High-water marks: High risk appetites? convex compensation, long horizons, and portfolio choice, *The Journal of Finance* 64, 1–36.

Teo, Melvyn, 2009, The geography of hedge funds, *Review of Financial Studies* 22, 3531–3561.

Appendix: Description Algorithm for Gross Returns

Here we describe the algorithm we use to compute monthly gross returns from reported monthly net-of-fee returns. Our approach allows for a monthly estimation of gross returns under different crystallization regimes (monthly or lower frequency).

The algorithm is based on the following set of assumptions:

1. The Gross Asset Value at the fund's inception (GAV_0) is equal to 100.

2. The algorithm is based on a single-investor assumption.

3. The management fee is paid monthly[14].

We start by defining the *unoberved* Gross Return at the end of month t ($GrossRet_t$):

$$GrossRet_t = \frac{GAV_t}{GAV_{t-1}} - 1 \tag{1}$$

where GAV_t and GAV_{t-1} are the unobserved Gross Asset Value at the end of month t and $t-1$, respectively.

The amount of Management Fee ($MgtFee_t$) paid in month t equals:

$$MgtFee_t = NAV_{t-1} \cdot (1 + GrossRet_t) \cdot \frac{MF\%}{12} \tag{2}$$

where $MF\%$ is the management fee (p.a.). The Total Management Fee Paid up to month t ($TotalMgtFeePaid_t$) is then:

$$TotalMgtFeePaid_t = \sum_{i=1}^{t} MgtFee_i \tag{3}$$

In addition to the management fee, we also calculate the amount of Interest Earned ($InterestEarned_t$) by the fund manager on excess cash and cash deposited in the margin account:

[14] This assumption can easily be relaxed to a different payment frequency by handling the payment of the management fee in the same way as the incentive fee. We nevertheless fix the payment frequency to monthly because an analysis of the managment fee is not the thrust of the analysis.

$$InterestEarned_t = NAV_{t-1} \cdot Rf_t \quad (4)$$

where Rf_t is the risk-free rate in month t. We take Interest Earned into account because CTAs typically hold up to 80% of the money in a cash account and earn interest on this cash. In the case of most other hedge fund strategies, this adjustment for Interest Earned is not required and can easily be omitted. Total Interest Earned on cash deposited ($TotalInterestEarned_t$) is the sum of all interest earned up to month t:

$$TotalInterestEarned_t = \sum_{i=1}^{t} InterestEarned_i \quad (5)$$

Using the above definitions, we define the Preliminary Net Asset Value at time t ($PrelNAV_t$) as:

$$PrelNAV_t = NAV_{t-1} \cdot (1 + GrossRet_t) - TotalMgtFeePaid_t - TotalIntEarned_t \quad (6)$$

As such, we subtract the management fee and the interest earned from the gross increase in NAV_{t-1}. Using $PrelNAV_t$ for the calculation of the incentive fee ensures that the manager only charges an incentive fee on performance *in excess of* any management fee charged and any risk-free return earned on cash. For the next set of equations, we introduce an indicator ($Cryst_t$) that takes on the value 1 in months where crystallization occurs, and zero otherwise.

The Accrued Incentive Fee ($AccrIncFee_t$) is a fraction of the performance – the incentive fee $IF\%$ – in excess of the current High-Water Mark (HWM_{t-1}):

$$\begin{cases} max(0, PrelNAV_t - HWM_{t-1}) \cdot IF\% & \text{if } Cryst_t = 0 \\ 0 & \text{if } Cryst_t = 1 \end{cases} \quad (7)$$

This means that, when no crystallization occurs, we only *accrue* the incentive fee. However, when crystallization does take place, the accrued incentive fee is paid to the fund manager. In that case we add any accrued incentive fee over the period since the last crystallization to the Incentive Fee Paid variable ($IncFeePaid_t$):

$$\begin{cases} IncFeePaid_{t-1} & \text{if } Cryst_t = 0 \\ IncFeePaid_{t-1} + max(0, PrelNAV_t - HWM_{t-1}) \cdot IF\% & \text{if } Cryst_t = 1 \end{cases} \tag{8}$$

At this point in time, the High-Water Mark (HWM_t) is also updated to the current Preliminary Net Asset Value if it exceeds the previous High-Water Mark:

$$\begin{cases} HWM_{t-1} & \text{if } Cryst_t = 0 \\ max(PrelNAV_t, HWM_{t-1}) & \text{if } Cryst_t = 1 \end{cases} \tag{9}$$

The Net Asset Value at time t (NAV_t) equals:

$$NAV_t = PrelNAV_t + TotalInterestEarned_t - IncFeePaid_t \tag{10}$$

Since no closed-form solution is available, we solve for the unobserved GAV_t numerically. In particular, we determine the value of GAV_t that equates the NAV_t computed in equation (10) – based on GAV_t – to the observed NAV at time t. We then store the obtained value of GAV_t and move to the next month, solving for GAV_t in an iterative way. When we charge fees in the main analysis, we also use the above equations to go from GAV_t to NAV_t.

Chapter 2

An Analysis of the Risk-Return Characteristics of Serially Correlated Managed Futures[1]

The Journal of Futures Markets
Vol. 36, No. 10, 992–1013 (2016)

2.1 Introduction

The historical track-record remains the most important piece of information in the evaluation of potential hedge fund managers. This is the case as information on the alpha-models used by the managers can only be inferred from their track-record. The models themselves remain strictly proprietary. As a consequence, past returns will remain a key element in manager selection. An important consideration in this regard, is the degree of persistence in managers' reported returns. If fund managers' returns exhibit persistence at certain frequencies, then manager selection based on past performance can potentially add value along this time series dimension.

In this article we provide empirical evidence that value can potentially be added through incorporating serial correlation patterns in Managed Futures' self-reported returns in the investment process. In particular, we find that Managed Futures funds that exhibit higher degrees of positive serial correlation – based on the unweighted sum of autocorrelations – exhibit distinctly different risk-return profiles and outperform funds that exhibit lower degrees of serial

[1]This chapter is based on joint work with Péter Erdős (RPM Risk & Portfolio Management) and John Sjödin (RPM Risk & Portfolio Management and Ghent University).

correlation. A portfolio of more positively autocorrelated Managed Futures funds displays higher risk-adjusted performance and lower drawdowns.

Application of multifactor models, including models using the recently proposed risk factors suggested by Baltas and Kosowski (2012) as well as the more commonly used hedge fund risk factors of Fung and Hsieh (2004), indicate a significantly positive risk-adjusted excess return ('alpha') of approximately 6 percent p.a. Interestingly, the models univocally suggest a lower explanatory power in the case of the more positively serially correlated Managed Futures funds. This finding of a low explanatory power of multifactor models coupled with risk-adjusted outperformance corroborates some recent findings in the literature on performance persistence in both the hedge fund and mutual fund performance literature.[2]

In particular, Sun, Wang, and Zheng (2012) propose a "Strategy Distinctiveness Index" (SDI) constructed as 1 minus the correlation between a hedge fund's historical returns and the returns of its peers. The objective of Sun, Wang, and Zheng their measure is to capture the degree to which hedge fund managers follow unique investment strategies. The authors find that higher strategy distinctiveness is associated with better future fund performance. Similarly, Titman and Tiu (2011) show that that hedge funds with lower R^2s with regard to systematic factors have higher Sharpe ratios, higher information ratios, and higher alphas. They conjecture that funds that have more confidence in their abilities will expose their investors less to factor risk.

Our results are consistent with the above findings. Sorting Managed Futures funds on the degree of serial correlation results in a subset of funds that outperform peers exhibiting lower degrees of serial correlation. Coincidentally, these more positively serially correlated funds' returns are found to be less well explained by existing multifactor models. This seems to suggest that the serial correlation we observe is a consequence of the unique investment strategies followed by these managers.

However, self-reported returns do not necessarily reflect all risks inherent to investing in hedge funds and thus might overstate the actual return experience of investors. Therefore, we explore several alternative explanations for the observed premium. Amongst others, we consider attrition rates and the associated delisting bias as well as exposure to tail risk as potential explanations for the observed outperformance. Despite slightly higher attrition rates among more positively serially correlated managers, we find that a potential delisting bias is unable to fully explain the observed outperformance.

The rest of this paper is structured as follows. The relevant literature is summarized and discussed in section 2.2. Section 2.3 describes the Managed Futures space considered for the analysis. In section 2.4 we outline the methodology used to determine the degree of persistence in Managed Futures funds' self-reported returns. We analyze the risk-return characteristics and potential drivers for the observed premium in section 2.5. Section 2.6 concludes.

[2] We thank an anonymous referee for calling attention to this connection with the recent literature.

2.2 Related Literature

Evidence of performance persistence among hedge funds is, of course, not new. Although early hedge fund literature gravitates towards a lack of performance persistence in hedge funds' self-reported returns (see inter alia ?Brown and Goetzmann, 2003; Capocci and Hübner, 2004; Malkiel and Saha, 2005), more recent contributions present evidence of performance persistence.

In particular, Agarwal and Naik (2000) find persistence at the monthly frequency,Baquero, ter Horst, and Verbeek (2005) find persistence at the quarterly level, and Agarwal, Daniel, and Naik (2009) and Kosowski, Naik, and Teo (2007) find evidence of persistence among funds at annual horizons. Regarding Managed Futures, Schneeweis, Spurgin, and McCarthy (1997) find, based on a limited set of CTAs, that there is some performance persistence and that multi-advisor Managed Futures funds display more persistence than single advisor CTAs. More recently, Gregoriou, Hübner, and Kooli (2010) find performance persistence over horizons of at least one quarter. A the same time, they note that most of this persistence disappears when evaluating managers' ability to remain within the top quartile of top performing funds.

There is, however, one potential complication that accompanies much of the observed performance persistence in hedge funds' returns. The observed predictability may, to a large extent, be driven by illiquidity in the funds' underlying positions. Getmansky, Lo, and Makarov (2004) show that illiquidity, caused by stale prices, can lead to spurious serial correlation in hedge funds' self-reported returns. The authors conclude that the performance persistence documented by Agarwal and Naik (2000) and others can be traced down to spurious serial correlation. These results are corroborated by Eling (2009) who, based on a review of the existing literature as well as new evidence, shows that illiquid hedge fund categories such as Arbitrage and Emerging Markets exhibit very high levels of performance persistence, while more liquid hedge fund strategies have low levels of persistence. Still, Kosowski, Naik, and Teo (2007) argue that some hedge funds in their sample continue to exhibit performance persistence at annual horizons, even after controlling for the impact of spurious serial correlation as detailed above.

Managed Futures funds' self-reported monthly returns, however, are a notable exception. Unlike most other hedge fund categories, Managed Futures funds' returns do not exhibit autocorrelation, on average.[3] This empirical finding is consistent with the particular nature of Managed Futures funds' strategies. These funds only trade highly liquid securities and are therefore very unlikely to exhibit positive autocorrelation due to illiquidity and smoothing.[4]

[3]In the case of Managed Futures and Dedicated Short Bias hedge funds, Getmansky, Lo, and Makarov (2004) obtain smoothing-parameter estimates that suggest that no unsmoothing of the returns is needed.

[4]This point is worth stressing, especially in light of recent evidence that performance predictability in equity hedge funds tends to weaken when taking into account liquidity risk (Brandon and Wang, 2013). Sadka (2010) finds that sorting Managed Futures into deciles based on their exposure to an (equity) liquidity risk factor does not yield a significant (Fung-Hsieh 7-factor) alpha. However, as Managed Futures do not trade individual equities, existing

In spite of the liquid nature of Managed Futures funds' strategies and the absence of high levels of serial correlation, Khandani and Lo (2011) nevertheless find evidence that, among the different hedge fund categories they consider, Managed Futures exhibit the largest 'illiquidity' premium. More specifically, the authors conclude that Managed Futures funds that exhibit higher degrees of positive autocorrelation outperform funds that exhibit lower degrees of positive autocorrelation. This finding is intriguing as this hedge fund category provides a special case where positive autocorrelation is unlikely to be driven by illiquidity. This suggests that there is *cross-sectional* variation in the degree of serial correlation in Managed Futures funds' returns that conveys information on future performance.

Two apparent hedge fund return profiles can be expected to yield persistence. First, we can imagine funds that exhibit highly persistent small positive returns. While such a return profile can be the result of consistently exploiting a mispricing, it can also be the result of a manager's decision to adopt a 'short-option' or 'short-volatility' profile. If the latter proves to be the case, one should see a breakdown in the profitability of these funds in periods of market stress. Second, we would also observe persistence in returns among managers that report return profiles that show occasional high positive return months, but many small negative months in between. In that case, the return behavior resembles a 'long-option' or 'long-volatility' profile. Such a profile carries a number of characteristics of CTAs' trend-following nature. For example, Fung and Hsieh (2001) make use of long-option strategies (lookback-straddles) to model the performance of trend following funds.

Furthermore, trend-following is a divergent risk-taking strategy (see Rzepczynski, 1999; Chung, Rosenberg, and Tomeo, 2004; Greyserman and Kaminski, 2014). That is, unlike convergent strategies where a manager will consider adding to an existing position when a perceived mispricing increases, trend-following approaches generally dictate closing positions when trends fail to materialize. This suggests that trend-followers can be expected to incur a lot of small losses, perhaps for extended periods of time, until market conditions allow clear trends to emerge. We attempt to determine the extent to which Managed Futures funds sorted on serial correlation exhibit one of the above-mentioned return profiles similar to being short- or long volatility and whether their performance breaks down in periods of market stress.

Our work is similar in spirit to the work of De Souza and Gokcan (2004), who propose using a measure of pure persistence, the Hurst exponent, to aid in hedge fund manager selection. The authors find that portfolios of hedge funds with a high Hurst exponent exhibit higher returns, lower standard deviations, and lower drawdowns. Unfortunately, their work does not cover Managed Futures.

Autocorrelation in Managed Futures funds' returns has been a topic of interest in recent empirical work. Burghardt and Liu (2013) demonstrate that trend-following Managed Futures exhibit negative autocorrelation over short

liquidity measures based on (individual) equities might prove unsatisfactory in analyzing a potential liquidity risk to which Managed Futures are exposed.

horizons of up to six months. The authors note that failing to account for negative autocorrelation in returns might yield biased performance statistics when scaling estimates of volatility. Another important question, which has not been addressed to the authors best knowledge, is the relationship between autocorrelation patterns in Managed Futures' returns and subsequent performance. Khandani and Lo (2011) their finding of a positive 'illiquidity' premium in Managed Futures seems to suggest a positive relationship. However, a more in-depth analysis is needed, as the autocorrelation patterns might in fact be indicative of specific risks taken by these managers. In what follows, we attempt to shed light on this matter.

2.3 Data

The data come from BarclayHedge. We rely on BarclayHedge as this is the most comprehensive database on Managed Futures that is available to researchers and practitioners. In addition, Joenväärä, Kosowski, and Tolonen (2012), in their comparison of five major publicly available hedge fund databases, find that BarclayHedge has the largest percentage of defunct funds (65%), thus making it least likely to suffer from survivorship bias. Following related literature, we only include the post-1994 period to avoid potential survivorship bias, as most databases only started collecting information on defunct funds from 1994 onwards.

We filter the dataset in several respects. First, we classify the Managed Futures programs in different categories based on the funds' self-reported strategy description.[5] In the process, we remove funds whose description indicates that they invest exclusively in options. If a fund reports multiple share classes for the same program, we only incorporate the fund's flagship program, which we identify as the share class with the longest track-record and highest assets-under-management (AUM). Second, we only include programs denominated in USD and EUR, and convert the EUR-denominated returns and AUM to USD using the end-of-month spot USD/EUR exchange rate. We remove funds with missing observations as well as zero-return observations at the start and end of a fund's track-record. To account for backfill bias, we also remove the first 12 observations of a fund's track-record (see, for example, Kosowski, Naik, and Teo (2007)). To ensure that our results apply to funds that can be considered part of the investable universe for investors, we remove funds whose returns exhibit unusually low levels of variation. To this end, we discard funds for which the standard deviation of the observed returns is lower than 2% p.a.

Similarly to Getmansky, Lo, and Makarov (2004) and Khandani and Lo (2011) we require a track-record of at least 5 years for a fund to be included. This minimum requirement on the track-record is needed to ensure a sufficient number of observations to be able to properly estimate the autocorrelation pattern in a fund's self-reported returns. Imposing this additional requirement,

[5]Despite the possibility of strategic self-misclassification, Brown and Goetzmann (2001) find that self-reported descriptions do almost as well as return-based procedures.

we obtain a dataset of 677 Managed Futures programs, 207 currently live and 470 that have stopped reporting ('defunct') as of the end of 2013. Summary statistics for the funds are reported in Table 2.1.

The statistics on the standard first order autocorrelation coefficient (ρ_1) corroborate the finding of no autocorrelation, on average, among Managed Futures (see Getmansky, Lo, and Makarov, 2004). Finally, the reported AUM indicate that our dataset covers US$ 157.2bn, as of the end of 2013.

2.4 Methodology

To measure the degree of autocorrelation in Managed Futures funds' returns, we calculate a fund's autocorrelation function based on the past five years of return data. Given the generally low levels of serial correlation in Managed Futures, we opt for an approach where we sum up the autocorrelation function up to lag 12, rather than focusing on the first order autocorrelation.[6] As such, our measure of serial correlation becomes,

$$\hat{P} = \sum_{i=1}^{12} \left(\hat{\rho}_{t-i} + \frac{1}{T-i} \right) \tag{2.1}$$

where $\hat{\rho}_{t-i}$ is the estimated autocorrelation at lag i and T is the sample size. Kendall and Stuart (1976) show that under the null hypothesis of serial independence, the $i-th$ sample autocorrelation is biased in small samples and has an expected value of $\frac{-1}{T-i}$. Therefore, our measure includes a small sample bias-correction whose importance is meaningful in this case as we have only 60 observations ($T = 60$).

Levich and Rizzo (1999)show that, in the case of small but persistent autocorrelation, the unweighted sum of autocorrelations has higher power in detecting persistence compared to conventional tests for autocorrelation such as the Durbin-Watson h and m tests, Bartlett-test, Box-Pierce Q-test, the LM test of Breusch (1978) and Godfrey (1978), and the variance ratio test. The environment for which these authors have developed their measures of persistence is very similar to the case of Managed Futures. Managed Futures, on average, do not exhibit significant autocorrelation, at least, based on conventional measures (see Table 2.1, panel A). However, this observation does not rule out very small, but persistent autocorrelation, which cannot be detected using conventional tests. Such a return characteristic can be an indication of superior managerial skills, in which case it is of considerable importance in portfolio selection.

Therefore, to be able to detect small, but persistent autocorrelation, our ranking relies on a measure that is almost identical to the one proposed by Levich and Rizzo (1999). The only difference is that we account for small sample bias. It is important to note that this way, we retain important information contained

[6] 12 months is consistent with the convention in the momentum literature and the presence of time-series momentum in futures markets (see Moskowitz, Ooi, and Pedersen, 2012).

Table 2.1: Managed Futures Data

Panel A: Summary Statistics

	Mean	Median	Standard deviation
Mean return	0.81%	0.61%	0.96%
Minimum return	-13.74%	-10.79%	11.38%
Maximum return	21.21%	15.32%	21.42%
Standard deviation	5.53%	4.41%	4.44%
Skewness	0.544	0.436	0.894
Kurtosis	5.778	4.357	4.401
Size ($US m)	208.43	24.39	1057.78
Age (Years)	9.8	8.3	4.3
ρ_1	0.02	0.01	0.13

Panel B: Evolution Data Set

Year	Live funds	Defunct	AUM ($US bn)
1999	308	24	28.93
2000	312	48	27.12
2001	317	71	33.57
2002	332	84	40.23
2003	346	103	68.58
2004	366	123	106.69
2005	395	144	96.77
2006	422	176	132.67
2007	448	194	161.59
2008	452	222	175.89
2009	426	251	180.60
2010	404	273	215.09
2011	367	310	217.75
2012	315	362	176.95
2013	207	470	157.23

Notes: this table reports summary statistics for the data set of Managed Futures. Panel A reports statistics on the monthly returns. Panel B reports end-of-year figures.

in the sign of ρ_{t-i} and do not downweight autocorrelation at higher lags which could result to miss out information related to performance persistence.

To provide some rationale how this measure works for detecting performance persistence, consider a white noise process. In this case, the probability that the first N autocorrelation coefficients are all positive is considerably lower than the probability that half the coefficients are positive and half are negative.[7] In such cases, the sum of autocorrelations might be more informative.

On a statistical ground, our measure is closely related to spectral measures. For example, let $f(0)$ denote the zero frequency spectrum of the returns. The spectral density of interest can then be given by

$$f(0) = \omega_{t-0} + 2\sum_{i=1}^{\infty} \omega_{t-i} \tag{2.2}$$

where ω stands for the autocovariance function. If we divide both sides of the equation by the variance of the returns,

$$f^{*}(0) = 1 + 2\sum_{i=1}^{\infty} \rho_{t-i} \tag{2.3}$$

that is, the normalized spectrum at frequency zero is the sum of autocorrelations (see among others Cochrane, 1988; Lo and MacKinlay, 1988). In applications the infinite sum on the right-hand side must be truncated. Indeed, we truncate the estimation and sum the unweighted autocorrelations up to lag 12. In this sense Eq. (2.1) is closely related to zero frequency spectrum estimators.[8]

It is quite straightforward that after correcting for the small-sample bias, if Managed Futures funds' returns are uncorrelated, Eq. (2.3) is equal to one and our measure (Eq. (2.1) equals to approximately zero. Under performance persistence, returns exhibit positive autocorrelation and Eq. (2.1) is above 0. Under long-term mean-reversion in Managed Futures funds' performance, returns are negatively serially correlated and P is negative.

[7] Assuming a white noise process and after correcting for small sample bias, the chance that half of the autocorrelations is positive is exactly 50%. As the number of positive autocorrelation is binomially distributed in the case of white noise, if 9 out of the 12 autocorrelation coefficients estimated to be positive, the null hypothesis of white noise can be rejected at conventional levels of significance, independently of the magnitude of autocorrelation coefficients.

[8] $\hat{P} = \frac{\hat{f}^{*}(0)-1}{2}$. Our estimation of $\hat{P}$ is matching the truncated uniform kernel-based estimation in Andrews (1991). If the truncated kernel is $x(i/k)$, $P = \sum_{i=1}^{\infty} x(i/k)\rho_{t-i}$, where $x = \begin{cases} 1 & \text{if } i/k \leq 1 \\ 0 & \text{otherwise} \end{cases}$. Moreover, White (1980) and Hansen (1982) also apply truncated and unweighted estimators to Eq. (2.3).

2.5 Results

2.5.1 Risk-Return Characteristics of Sorted Portfolios

Using our measure of return persistence, we rank our sample of Managed Futures funds and divide them into quintile portfolios, with the highest (lowest) quintile portfolio consisting of those funds with the highest (lowest) degree of persistence measured by the unweighted sum of the first 12 autocorrelation coefficients, as in Levich and Rizzo (1999). We update the ranking of the funds at the end of every month, effectively rebalancing the portfolio on a monthly basis. For the purpose of the analysis, we construct quintile portfolios both on an equal-weighted and asset-weighted basis (using the funds' reported AUM at $t-1$.[9] To avoid the portfolio construction suffering from look-ahead bias, we insert a zero-return the first month after a fund stops reporting.[10]

We report results in Table 2.2 for the quintiles of interest. The inception date of the portfolios is 1999, as we require a 5-year burn-in period to estimate the autocorrelation structure for the set of funds. Absolute performance, measured using the compound annual growth rate (CAGR), suggests that more positively autocorrelated Managed Futures funds' (Q5) outperformed their less positively autocorrelated peers (Q1) on an absolute return basis over the 1999-2013 period. The upper quintile portfolio of most positively autocorrelated Managed Futures posted a CAGR of 7.38% p.a., compared to 4.52% p.a. for the lower quintile portfolio. This result is in line with Khandani and Lo (2011) their earlier finding of the presence of an illiquidity premium in Managed Futures.

Sorting managers based on serial correlation thus appears to yield portfolios with higher raw performance. p-values for a standard difference in means test, based on a bootstrap with a 1000 replications, however, suggests that the mean average returns are not significantly different at conventional levels, with a p-value of 0.16. Average monthly performance, of course, does not consider the level of risk taken.

Higher average returns are consistent with the argument that, as positive serial correlation is commonly considered a measure of illiquidity (see Getmansky, Lo, and Makarov, 2004) and, thus, illiquidity risk, positively autocorrelated returns may indicate higher risk. The general absence of illiquidity in Managed Futures funds' underlying positions makes this finding unexpected. Still, the higher expected returns may be a compensation for higher risk of some sort. If this is the case, we expect the top quintile portfolio (Q5) to exhibit higher levels of riskiness than the bottom quintile portfolio (Q1).

[9] Since small funds are generally not considered for investment, we perform a robustness check where we impose the additional requirement that the fund should have at least US$10 million AUM at rebalancing. Results are robust to such an AUM-based filter. Results available upon request.

[10] In this case, the information that a fund has stopped reporting in the following month is not available to an allocator at the time of rebalancing. As such, to avoid look-ahead bias, we should assume a certain allocation to that fund, even though the actual return is not observed. Later on we relax this arbitrary zero return assumption further, to account for the bias that voluntary reporting might induce.

To analyze is conjecture, we also report several measures of risk and risk-adjusted performance in Table 2.2. In particular, we report the monthly standard deviation, the (autocorrelation-adjusted) Sharpe ratio[11], maximum drawdown, and the Sortino ratio. Since controlling downside risk plays an important role in hedge funds and Managed Futures funds in particular, measures based on Lower Partial Moments (see Harlow and Rao, 1989) are also considered. The Sortino ratio (Sortino and Van Der Meer, 1991) is one commonly used measure of downside risk. We report this metric with a target return of zero. Finally, Maximum drawdown (MDD) is reported as this is a metric of particular relevance for practitioners in the Managed Futures industry.

The risk-adjusted performance measures indicate that the upper quintile portfolio outperforms the lower quintile, regardless of the particular risk measure used. Interestingly, the outperformance of the top quintile portfolio (Q5) seems to be mainly driven by lower volatility. As such, the Sharpe- and Sortino ratio are considerably higher[12] for the top quintile portfolio. Using the hypothesis testing methodology suggested by Ledoit and Wolf (2008) (henceforth, *LW*) we test whether the difference in Sharpe ratios for the top and bottom quintile is actually significantly different or not. We find this to be the case, as the difference is significant at conventional levels (p-value of 0.0062 and 0.08 for the AUM-weighted and equal-weighted portfolios, respectively).

This better risk-adjusted performance in terms of reward-to-variability is particularly important in Managed Futures space, as funds' programs are typically leveraged multiple times to obtain a certain target-level volatility. Maximum drawdown statistics indicate that a portfolio consisting of the most positively serially correlated funds exhibits drawdowns notably lower than that of the other portfolio. This finding suggests that the positive autocorrelation in Managed Futures, at least at first sight, does not lead to deeper drawdowns. The analysis so far yields a set of Managed Futures managers that outperform their peers. We should nevertheless first consider real-life limitations to investing in hedge funds before we can proceed.

Share restrictions such as the lockup period, advance notice period and the redemption frequency can limit an allocator's ability to exploit short-term persistence present in hedge funds[13]. However, compared to other hedge fund

[11] Annualized Sharpe ratios are adjusted for autocorrelation as suggested by Lo (2002). In particular, the reported Sharpe ratios are calculated as $SR(q) = \eta(q) \cdot SR$ with

$$\eta(q) \equiv \frac{q}{\sqrt{q + 2 \cdot \sum_{k=1}^{q-1} (q-k) \cdot \rho_k}} ,$$

Where SR is the regular Sharpe ratio on a monthly basis, is ρ_k is the $k-th$ order autocorrelation. $SR \cdot \eta(q)$ is then the annualized autocorrelation adjusted Sharpe ratio with $q = 12$.

[12] In unreported results, we find that failing to adjust the Sharpe ratio has a material impact as it increases (lowers) the ratio for the top (bottom) quintile portfolios, when compared to the adjusted Sharpe ratio. This is because the quintile portfolios themselves also exhibit positive (resp. negative) autocorrelation.

[13] Lockup refers to the initial amount of time investors are prohibited from withdrawing their investment. Once this lockup period is over, investors are allowed to withdraw their

categories, share restrictions are less stringent in the case of Managed Futures. One likely explanation for the lower restrictions is that redemptions are less costly for Managed Futures, as liquidity in futures markets makes these funds better able to scale down positions to meet redemptions. To illustrate this feature of Managed Futures, we report summary statistics on share restrictions for both Managed Futures and a composite of the other hedge fund categories that report to BarclayHedge. As lock-ups are uncommon for most hedge fund categories, with 70% of the funds having no lock-up restriction in place, we focus on advance-notice periods and redemption frequencies. In order to draw conclusions from the advance-notice period and redemption frequency, we need to analyze both in conjunction. Consider for example a fund that imposes for a one-day advance-notice period but nevertheless allows redemptions only quarterly. In that case, although the advance-notice is one day, redemption can take up to three months.

While a wide range of combinations is possible, the actual number of combinations is more limited in practice. For parsimony, we report in Table 2.3 the frequencies with which different combinations of share restrictions prevail, considering 40 combinations (based on 5 advance-notice bins and 8 redemption frequency bins).

Results in Table 2.3 illustrate that share restrictions are much less common for Managed Futures than for the other hedge fund categories. In particular, the vast majority of Managed Futures allow investors to redeem considerably more easily. Managed Futures generally allow redemption within the month, whereas far less the case for hedge funds.

But even if share restrictions are unrestrictive, considerable turnover required in maintaining the portfolios might still make implementation unrealistic. To investigate the turnover required, we report the change in the composition of the portfolios from month-to-month. We find that, while turnover is non-negligible, it is lowest for the upper quintile portfolio, at 12.7% per month. The lower quintile suggests a slightly higher turnover rate of 16.2%. The low turnover for both portfolios is to some extent the result of the fairly long track-record used in estimating the autocorrelation function, causing the resulting levels of autocorrelation to be fairly persistent. This suggests that this approach that relies on autocorrelation might have value in practice, especially in manager selection.

2.5.2 Performance Evaluation

The results above indicate that portfolios of Managed Futures funds based on serial correlation exhibit distinctly different risk-return characteristics. Now make use of a multifactor approach to try and identify the potential drivers of the observed outperformance. In particular, the standard approach in this context consists of assessing whether particular factors explain the performance of the different quintile portfolios.

capital only at pre-specified times of the year (dictated by the redemption frequency), and an advance notice is required for withdrawal.

Table 2.2: Summary Statistics Sorted Portfolios

	Mean Monthly Return	Standard Deviation	Sortino Ratio	Sharpe Ratio	MDD	CAGR
	Value-weighted Portfolios					
Q1 (low)	0.41%	2.83%	0.84	0.55	-16.88%	4.52%
Q5 (high)	0.61%	1.56%	2.88	1.19	-5.90%	7.38%
difference in means (p-val)	0.16		LW-statistic	3.298***		
BarclayHedge	0.55%	2.28%	1.53	0.98	-7.77%	6.40%
	Equal-weighted Portfolios					
	Mean Monthly Return	Standard Deviation	Sortino Ratio	Sharpe Ratio	MDD	CAGR
Q1 (low)	0.55%	2.67%	1.32	0.86	-10.69%	6.32%
Q5 (high)	0.49%	1.61%	2.25	1.15	-5.75%	5.88%
difference in means (p-val)	0.36		LW-statistic	1.694*		
BarclayHedge	0.49%	2.42%	1.31	0.81	-9.27%	5.69%

Notes: this table reports summary statistics on portfolios sorted portfolio exhibiting the highest degree of positive (negative) autocorrelation. The table reports the mean monthly return, the standard deviation of mean monthly returns, the annual Sortino ratio, the annual Sharpe ratio, maximum drawdown (MDD), and the compound annual growth rate (CAGR). A difference in means test, using a bootstrap with a 1000 replications is used to test the difference in average returns. The Ledoit-Wolf (LW) statistic tests the statistical significance of the difference in Sharpe ratios. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.3: Share Restrictions in Hedge Funds and Managed Futures

Panel A: Hedge Funds

		Redemption Frequency							
		Daily	Weekly	Bi-weekly	Monthly	Bi-monthly	Quarterly	Semi-annual	Annual
Advance Notice Period (days)	0	2.33%	1.04%	0.01%	1.30%	0.01%	0.78%	0.11%	0.09%
	1-31	2.62%	2.23%	0.20%	26.28%	0.30%	8.49%	0.56%	0.46%
	32-91	0.10%	0.09%	0.00%	13.33%	0.01%	19.91%	1.46%	1.76%
	92-180	0.00%	0.03%	0.01%	5.08%	0.00%	8.44%	1.00%	1.76%
	> 180	0.00%	0.00%	0.00%	0.01%	0.00%	0.08%	0.03%	0.08%

Panel B: Managed Futures

		Redemption Frequency							
		Daily	Weekly	Bi-weekly	Monthly	Bi-monthly	Quarterly	Semi-annual	Annual
Advance Notice Period (days)	0	11.72%	1.56%	0.00%	14.84%	0.00%	0.00%	0.00%	0.00%
	1-31	7.03%	7.03%	0.78%	44.53%	1.56%	0.78%	0.00%	0.00%
	32-91	0.00%	0.00%	0.00%	5.47%	0.00%	1.56%	0.00%	0.00%
	92-180	0.00%	0.00%	0.00%	3.13%	0.00%	0.00%	0.00%	0.00%
	> 180	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Notes: this table reports summary statistics on the share restrictions for Managed Futures and hedge funds. Results indicate the frequency with different combinations of advance notice and redemption frequency are employed.

While Managed Futures' dynamic nature of their strategies makes it difficult to model their returns, recent advances on (time-series) momentum in futures markets by Moskowitz, Ooi, and Pedersen (2012) and Baltas and Kosowski (2012) have led to an improved understanding of Managed Futures. Moskowitz, Ooi, and Pedersen (2012) present evidence that futures contracts' own past returns predict future returns. To exploit this predictability, the authors implement synthetic trading strategies that take both long- and short positions in a wide set of futures contracts, using information inferred from the contracts' (12-month) past returns. Their results also suggest these momentum factors capture the performance of Managed Futures returns and perform better than of the primitive trend-following strategy metrics (PTFS), suggested by Fung and Hsieh (2001). Baltas and Kosowski (2012) extend Moskowitz, Ooi, and Pedersen (2012) their approach and construct time-series momentum factors over different trading horizons. They show that a combination of these factors and the seven factors of Fung and Hsieh (2004) considerably improves the explanatory power of the model applied to Managed Futures' returns.

We incorporate these recent advances on performance evaluation to analyse the different quintile portfolios. In particular, we retrieve the data for Fung and Hsieh's 7-factor model and Baltas and Kosowski (2012) their momentum factors.[14] We then estimate multifactor models for the relevant value-weighted quintile portfolios for the 1999-2013 period for which all data is available. Results are reported in Table 2.4.

Examining the observed variance explained across models, using the adjusted-R^2, we find that more positively autocorrelated Managed Futures' returns are less well explained, both in the case of the momentum factors and a combination of the momentum factors and Fung and Hsieh's 7-factor model. The upper quintile portfolio displays considerably lower loadings on the different momentum factors, although the momentum factors remain significant at conventional levels. Looking at the upper quintile's risk-adjusted performance, we find that it is the only portfolio that exhibits a statistically and economically significant positive alpha (approximately 0.49% per month, or 6% p.a.). Nevertheless, the models' low explanatory power suggest that these programs are employing truly different strategies than most Managed Futures.[15] The lack of statistical significance of the factors proposed by Fung and Hsieh (2004) further suggest that these funds are not loading on any of the other risk-factors commonly associated with other hedge fund categories. This result is in accordance of the findings of Sun, Wang, and Zheng (2012) who show that hedge fund managers who produce

[14]The momentum factors are made available by Baltas and Kosowski (2012) at `http://www3.imperial.ac.uk/riskmanagementlaboratory/risklabsections/centreforhedgefundsresearch/baltas_kosowski_factors`. Data for the PTFS-factors are retrieved from the David Hsieh's home page `http://faculty.fuqua.duke.edu/~dah7/DataLibrary/TF-FAC.xls`.

[15]In unreported tests, we also analyse whether liquidity risk, proxied using a tradable (equity) liquidity factor of Pastor and Stambaugh (2003) their measure of illiquidity (available on Robert F. Stambaugh's home page `http://finance.wharton.upenn.edu/~stambaugh/`) sheds additional light on the outperformance. However, the risk-factor is not statistically significant at conventional levels. Results available upon request.

Table 2.4: Multifactor Model - Momentum Factors and Fung and Hsieh (2004) Factors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Q1	Q5	Q5-Q1	BarclayHedge	Q1	Q5	Q5-Q1	BarclayHedge
MOMM	0.288***	0.0669**	-0.221***	0.239***	0.305***	0.0697**	-0.235***	0.249***
	(0.0599)	(0.0333)	(0.0621)	(0.0373)	(0.0673)	(0.0314)	(0.0693)	(0.0393)
MOMW	0.179**	0.0909**	-0.0885	0.195***	0.167**	0.0816*	-0.0851	0.180***
	(0.0750)	(0.0445)	(0.0739)	(0.0521)	(0.0808)	(0.0458)	(0.0800)	(0.0564)
MOMD	0.0539	0.0378	-0.0161	0.0984**	0.0171	0.0117	-0.00543	0.0606
	(0.0610)	(0.0430)	(0.0606)	(0.0449)	(0.0631)	(0.0423)	(0.0699)	(0.0460)
S&P 500					0.0364	0.0399	0.00358	0.0169
					(0.0508)	(0.0295)	(0.0540)	(0.0379)
SCMLC					0.0276	0.0379	0.0103	0.0456
					(0.0997)	(0.0331)	(0.0956)	(0.0582)
10Y					-0.143**	-0.0447	0.0980	-0.122***
					(0.0578)	(0.0315)	(0.0671)	(0.0379)
CREDITSPR					0.163**	0.0426	-0.121	0.121**
					(0.0776)	(0.0428)	(0.0908)	(0.0531)
PTFSCOM					0.00195	-0.00292	-0.00487	-0.000548
					(0.0158)	(0.0106)	(0.0163)	(0.00970)
PTFSFX					0.0235*	0.0188**	-0.00477	0.0190*
					(0.0133)	(0.00828)	(0.0130)	(0.00995)
PTFSBD					0.0295**	0.0131	-0.0164	0.0275***
					(0.0135)	(0.00923)	(0.0142)	(0.0105)
Constant	-0.00133	0.00485***	0.00617***	0.000332	-0.000951	0.00499***	0.00594***	0.000758
	(0.00211)	(0.00126)	(0.00202)	(0.00157)	(0.00225)	(0.00134)	(0.00220)	(0.00163)
Observations	157	157	157	157	157	157	157	157
Adj. R^2	0.285	0.130	0.157	0.416	0.355	0.210	0.183	0.505

Notes: the table analyzes the monthly returns of the different quintile portfolios using Baltas and Kosowski (2012) their momentum factors and a combination of Baltas and Kosowski (2012) their factors and Fung and Hsieh (2004) their 7-factor model. The Fung and Hsieh (2004) factors are the Standard & Poors 500 index monthly total return (S&P 500); the spread return between Russell 2000 index monthly total return and Standard & Poors 500 monthly total return (SCMLC); The monthly change in the 10-year treasury (constant maturity) yield (10Y); the monthly change in the Moody's Baa yield less 10-year treasury constant maturity yield (CREDIT SPR); Fung and Hsieh (2001) their Bond Trend-Following Factor (PTFSBD), Currency Trend-Following Factor (PTFSPX), and Commodity Trend-Following Factor (PTFSCOM).
Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

returns less explainable by factors are more likely to possess managerial skills as they pursue more distinct strategies.

2.5.3 Alternative Explanations for the Premium

While so far our analysis yields notable results with regard to the risk-adjusted performance of more positively autocorrelated Managed Futures funds, it is instructive to explore alternative explanations that might explain the observed premium. To this end, we examine whether reliance on particular strategies, possible differential performance during adverse market states, attrition rates, and backfill bias might explain the performance.

2.5.3.1 Relationship with Managed Futures' Strategies and Funds' Traits

The portfolios' composition could be concentrated in Managed Futures categories that execute distinctly different strategies. In particular, funds could engage in trading strategies such as option writing, which might lead to different risk/return-profiles compared to the more dominant trend-following strategy. Non-trend-following strategies might therefore generate steady positive returns that induce positive serial correlation, but which might be followed by large losses. As mentioned in the data description, we have removed funds that indicate that they rely exclusively on option strategies.

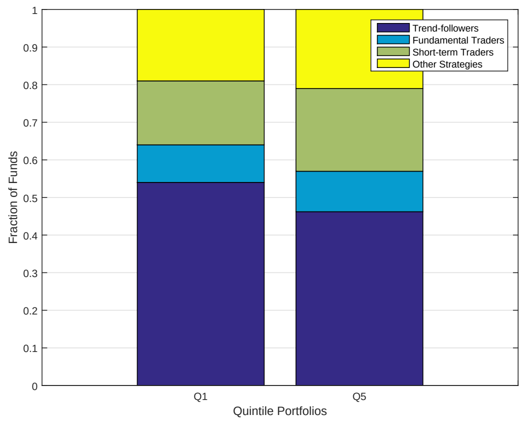
Nevertheless, it is instructive to report the composition of the quintiles of interest in terms of the strategies employed by the constituents. To this end, we employ the classification performed during the data handling. The results are reported in Figure 2.1.

The bar charts indicate that, while a portfolio consisting of positively autocorrelated Managed Futures seems to contain somewhat fewer (systematic) trend-followers, there are nevertheless no pronounced differences in the strategies employed by the managers included within every quintile portfolio. This suggests that the positive autocorrelation is not a feature of a particular strategy, but rather a feature of certain funds across different strategies.

There is a second dimension along which the strategies the funds follow might lead to a stronger performance of the upper quintile portfolio, compared to the other quintiles. In particular, differences in risk-adjusted performance might to some extent be driven by diversification gains. To analyze whether the potential for diversification gains differs across the different quintiles, we report the average pairwise correlation among the constituents prior to portfolio formation. We estimate pairwise correlations using the 5-year lookback window used to estimate the autocorrelation structure.

The results indicate that average pairwise correlation between any two funds is indeed lower in the case of the upper quintile portfolio. In particular, the pairwise correlation equals 0.11 for the upper quintile compared to 0.2 for the lowest quintile. This finding indicates that part of the strong performance is due to diversification gains. However, it also corroborates our earlier conjecture

Figure 2.1: Strategy Composition Quintile Portfolios



Notes: the stacked-bars report the composition of the different quintile portfolios over the sample period.

that these managers do not cluster around a particular investment approach. Instead, managerial skills might explain the good performance and low pairwise correlation with other Managed Futures.

Next to the strategies, we also analyse the average size and age of the funds included within every quintile. Given recent evidence that hedge fund performance is related to age and size (see Boyson (2008)), it is possible that the upper quintile consists of smaller or younger funds. The results for the average fund size suggest no differences in average fund size. The average fund size is USD 361m and USD 325m for the lower and upper quintile, respectively. A conventional t-test allows us to conclude that there are indeed no significant differences in the average size of funds in the extreme quintiles (p-value of 0.3265). In unreported results, we also observe that there are no significant differences in the age of the funds across quintiles.

2.5.3.2 Tail Risk

Of course, it is possible that there is a difference between what fund managers say they do, and what they actually do. Therefore we also consider an alternative approach to determine whether more positively autocorrelated Managed Futures take on tail risk. One manifestation of differential risk-taking should be evident when comparing performance during adverse market states. Fung and Hsieh (1997) are the first to use such an approach and show that Managed Futures exhibit a straddle-like pay-off. This feature of Managed Futures has been coined 'crisis alpha' by Kaminski and Mende (2011). Good overall performance of a portfolio investing in more positively autocorrelated Managed Futures might come at the expense of crisis alpha, i.e. strong performance during crisis times. Positively correlated funds' performance might break down during adverse market states and thus hamper their diversification benefit in a portfolio context.

We investigate the portfolios' performance during different market states, following the approach of Fung and Hsieh (1997). In particular, we group monthly returns of MSCI World Gross Total Return into five market states, ranging from sharp selloffs to rallies, by ranking the monthly gross returns. We then report the average performance of both the equity index and the portfolios of Managed Futures in the same period. For comparison, we perform a volatility adjustment such that the Managed Futures portfolios, ex-post, exhibit the same degree of volatility as the equity index. We do the adjustment in the following way

$$R_p^{adj} = \frac{\hat{\sigma}(R_{world} - R_f)}{\hat{\sigma}(R_p - R_f)} \cdot (R_p - R_f) + R_f \tag{2.4}$$

where $\hat{\sigma}()$ stands for the estimated standard deviation. R_{world} is the monthly gross return on the MSCI World Index, R_f is the monthly risk-free rate and R_p is the monthly return of the portfolio whose volatility we wish to scale. Since it is not possible to lever the interest rate component (proxied here by the risk-free rate) inherent in Managed Futures' returns, we subtract the risk-free rate from R_p when performing the volatility adjustment and then add it again afterwards.[16] The results are reported in Figure 2.2.

The results suggest that the higher performance of more positively autocorrelated Managed Futures does not lead to a deterioration of performance during adverse market states.

Another approach to analyzing whether Managed Futures funds in the top or bottom quintile are exposed to tail risk can be done using a regression approach. As described in the introduction, a likely explanation as to why we might expect persistence in the returns of Managed Futures has to do with the observation that their payoff resembles long volatility. To analyze whether the quintile portfolios of interest exhibit behavior similar to that of a put-option writing strategy, we proxy the performance of such a strategy using monthly returns on the CBOE S&P 500 PutWrite Index. Table 2.5 reports the results when we include this additional risk factor.

The outperformance of the upper quintile does not seem to be the result of taking on tail risk by engaging in (short) put-option writing on the S&P 500. In addition, the results on the long/short portfolio indicate that the upper and bottom quintiles' exposure with regard to this risk factor does not differ significantly. Interestingly, the BarclayHedge index appears to load positively on this risk factor, even after inclusion of the Fung and Hsieh (2004) factors.

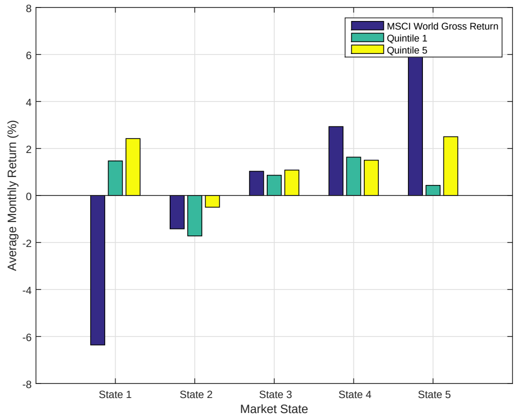
[16] While a Managed Futures program can be levered several times by changing the amount of margin held, this is not the case for the return earned on the cash held (i.e. risk-free rate). One should therefore subtract this return imbedded in a Managed Futures program's reported return when adjusting the volatility of a program.

Table 2.5: Multifactor Model - Portfolio Returns and Option Writing

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Q1	Q5	Q5-Q1	BarclayHedge	Q1	Q5	Q5-Q1	BarclayHedge
MOMM	0.290***	0.0688**	-0.222***	0.240***	0.305***	0.0704**	-0.235***	0.250***
	(0.0606)	(0.0322)	(0.0625)	(0.0374)	(0.0681)	(0.0315)	(0.0694)	(0.0382)
MOMW	0.209***	0.116***	-0.0928	0.218***	0.183**	0.0982**	-0.0848	0.205***
	(0.0721)	(0.0411)	(0.0740)	(0.0498)	(0.0866)	(0.0477)	(0.0842)	(0.0597)
MOMD	0.0800	0.0600	-0.0200	0.119**	0.0455	0.0405	-0.00496	0.105**
	(0.0623)	(0.0441)	(0.0625)	(0.0477)	(0.0648)	(0.0434)	(0.0705)	(0.0484)
S&P 500					-0.0314	-0.0290	0.00247	-0.0891
					(0.0938)	(0.0544)	(0.0873)	(0.0695)
SCMLC					0.0159	0.0259	0.0101	0.0273
					(0.0989)	(0.0342)	(0.0956)	(0.0562)
10Y					-0.137**	-0.0391	0.0981	-0.113***
					(0.0583)	(0.0326)	(0.0679)	(0.0397)
CREDITSPR					0.150*	0.0289	-0.121	0.0997*
					(0.0793)	(0.0463)	(0.0942)	(0.0557)
PTFSCOM					-8.74e-05	-0.00499	-0.00490	-0.00373
					(0.0158)	(0.0109)	(0.0162)	(0.00999)
PTFSFX					0.0220	0.0172**	-0.00480	0.0166*
					(0.0133)	(0.00838)	(0.0130)	(0.00984)
PTFSBD					0.0282**	0.0118	-0.0164	0.0255**
					(0.0137)	(0.00939)	(0.0145)	(0.0102)
PUTWRITE	0.0987*	0.0840**	-0.0146	0.0793*	0.126	0.128	0.00206	0.198**
	(0.0584)	(0.0365)	(0.0600)	(0.0446)	(0.125)	(0.0780)	(0.122)	(0.0924)
Constant	-0.00240	0.00393***	0.00633***	-0.000533	-0.00183	0.00409***	0.00593***	-0.000624
	(0.00223)	(0.00128)	(0.00219)	(0.00160)	(0.00230)	(0.00131)	(0.00222)	(0.00172)
Observations	157	157	157	157	157	157	157	157
Adj. R^2	0.295	0.155	0.157	0.426	0.359	0.225	0.183	0.520

Notes: the table analyzes the monthly returns of the different quintile portfolios using Baltas and Kosowski (2012) their momentum factors and a combination of Baltas and Kosowski (2012) their factors and Fung and Hsieh (2004) their 7-factor model. The Fung and Hsieh (2004) factors are the Standard & Poors 500 index monthly total return (S&P 500); the spread return between Russell 2000 index monthly total return and Standard & Poors 500 monthly total return (SCMLC); The monthly change in the 10-year treasury (constant maturity) yield (10Y); the monthly change in the Moody's Baa yield less 10-year treasury constant maturity yield (CREDIT SPR); Fung and Hsieh (2001) their Bond Trend-Following Factor (PTFSBD), Currency Trend-Following Factor (PTFSPX), and Commodity Trend-Following Factor (PTFSCOM). Finally, an option strategy involving writing out-of-the-money put options on the S&P 500 is captured using CBOE PutWrite index (PUTWRITE).
Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Figure 2.2: Performance During Different Market States



Notes: the bar chart reports the average monthly return during different market states. Market states are identified by ranking monthly gross returns of the MSCI World into 5 different quintiles. Average (volatility-adjusted) monthly returns for the quintile portfolios during the corresponding months are reported.

2.5.3.3 Attrition and Delisting Bias

While differences in risk-taking might not be evident from the trading strategies employed or performance during adverse market states, such differences may nevertheless show up when examining the funds' attrition rates. Attrition rates allow us to quantify potential risks not captured by the funds' self-reported returns. hedge funds in general and Managed Futures in particular have high attrition rates, as is evident from Table 2.1. Arnold (2013) notes that while attrition of Managed Futures is high, real failures are considerably lower, suggesting that many liquidations may not be damaging to investors. Nevertheless, given the voluntary nature of hedge fund databases, managers might fail to report further losses to the investors by not reporting last months' performance. Consequently, returns might not reflect the actual losses of investors. The delisting bias that such behaviour induces, has been analysed in context of hedge fund databases. Edelman, Fung, and Hsieh (2013) conclude that missing returns of successful funds tend to offset the delisting bias in the missing returns of liquidating funds.

Nevertheless, we analyse attrition rates and the possible impact of backfill bias on our results. We start by counting the number of fund delistings that occur for every quintile portfolio in the period immediately after rebalancing. In particular, we count the number of instances where our portfolio construction would have invested in funds that no longer report in the subsequent period. This provides a first useful proxy of risks that do not show up in the funds' self-reported returns.

We find that the fraction of delistings is slightly higher in the case of the upper quintile portfolio with an attrition rate of 26% (108 delistings), compared to 22% (88 delistings) in the case of the lower quintile. These results suggest that there are more fund failures among positively autocorrelated Managed Futures, although subdued. Nevertheless, this result does suggest that perhaps the outperformance is be driven by delisting bias.

Therefore, we attempt to explicitly correct for the delisting bias. In particular, we repeat the portfolio approach outlined above, but assume a -4.5% return in the first month that the fund fails to report to the database. This -4.5% return corresponds to the average compounded omitted return for the Lipper TASS and HFR database found by Jorion and Schwarz (2013). Correcting for delisting bias in this way takes into account the higher incidence of fund delistings in certain quintile portfolios. This is necessary as the likelihood of a fund becoming delisted seems to be positively correlated to higher degrees of positive autocorrelation in the programs' returns. The results for the value-weighted quintile portfolios are reported in Table 2.6.

We find that the performance of positively autocorrelated Managed Futures seems to persist, even when we correct for delisting bias using a conservative -4.5% return. This is particularly the case for the AUM-weighted portfolios, but appears to be the less the case for the equal-weighted portfolios.

2.6 Conclusion

In this paper, we developed and applied a measure for detecting low but persistent levels of performance persistence in hedge funds' self-reported returns. We applied this measure to Managed Futures, a hedge fund category that is unlikely to exhibit spurious serial correlation due to smoothing and illiquidity in underlying positions.

We make several contributions to the existing literature on autocorrelation patterns in hedge funds and Managed Futures in particular. First, we corroborate earlier findings in that we provide additional evidence of the existence of a premium in Managed Futures, using an alternative hedge fund database. Second, using a multifactor analysis, we find that the observed outperformance of funds sorted on the degree of persistence in their returns cannot be explained using existing models. This suggests that the returns generated by these funds are distinctly different. Third, we show that the premium is unlikely to be explained by a reliance on particular strategies, fund size, a compensation for tail risk, attrition rates, and delisting bias. Given considerably lower share restrictions for Managed Futures, our results suggest that incorporating serial correlation may improve the manager selection and allocation process.

The above results suggest that the observed persistence might be a proxy of fund skills. If a fund manager has a good trading approach that fits the prevailing market environment at a given period in time, that fund is expected to persistently generate gains. Of course, a particular trading approach should not be expected to work indefinitely since the market environment regularly

Table 2.6: Results Correction for Delisting Bias

	Value-weighted Portfolios					
	Mean Monthly Return	Standard Deviation	Sortino Ratio	Sharpe Ratio	MDD	CAGR
Q1 (low)	0.40%	2.83%	0.81	0.53	-17.14%	4.39%
Q5 (high)	0.57%	1.59%	2.56	1.12	-6.61%	6.91%
Difference in means (p-val)	0.19		LW-statistic	2.94**		
BarclayHedge	0.55%	2.28%	1.53	0.98	-7.77%	6.40%

	Equal-weighted Portfolios					
	Mean Monthly Return	Standard Deviation	Sortino Ratio	Sharpe Ratio	MDD	CAGR
Q1 (low)	0.50%	2.68%	1.18	0.78	-10.93%	5.71%
Q5 (high)	0.43%	1.62%	1.88	0.96	-8.22%	5.09%
Difference in means (p-val)	0.38		LW-statistic	1.34		
BarclayHedge	0.43%	1.62%	1.53	0.98	-7.77%	6.40%

Notes: this table reports the results for a robustness check where we repeat the portfolio construction, but at the same time impose a hypothetical -4.5% return in the first month a fund stops reporting to Barclayhedge. The table reports the mean monthly return, the standard deviation of mean monthly returns, the annual Sortino ratio, the annual Sharpe ratio, maximum drawdown (MDD), and the compound annual growth rate (CAGR).A difference in means test, using a bootstrap with a 1000 replications is used to test the difference in average returns. The Ledoit-Wolf (LW) statistic tests the statistical significance of the difference in Sharpe ratios. *** p<0.01, ** p<0.05, * p<0.1

changes. As such, rebalancing the portfolio is required. Finally, we note that our results suggest that, while it is unlikely that the outperformance of more positively autocorrelated Managed Futures funds is driven by delisting bias, slightly higher attrition rates require close monitoring and risk management.

Bibliography

Agarwal, Vikas, Naveen D. Daniel, and Narayan Y. Naik, 2009, Role of managerial incentives and discretion in hedge fund performance, *The Journal of Finance* 64, 2221–2256.

Agarwal, Vikas, and Narayan Y. Naik, 2000, Multi-period performance persistence analysis of hedge funds, *Journal of Financial and Quantitative Analysis* 35, 327–342.

Andrews, Donald WK, 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica: Journal of the Econometric Society* 817–858.

Arnold, Julia, 2013, Survival of commodity trading advisors: Systematic vs. discretionary ctas., *Unpublished working paper, Centre for Hedge Fund Research, Risk Management Laboratory, Imperial Colledge Business School* .

Baltas, Akindynos-Nikolaos, and Robert Kosowski, 2012, Momentum strategies in futures markets and trend-following funds, *Unpublished working paper, Centre for Hedge Fund Research, Risk Management Laboratory, Imperial Colledge Business School* .

Baquero, Guillermo, Jenke ter Horst, and Marno Verbeek, 2005, Survival, look-ahead bias, and persistence in hedge fund performance, *Journal of Financial and Quantitative Analysis* 40, 493–517.

Boyson, Nicole M, 2008, Hedge fund performance persistence: a new approach, *Financial Analysts Journal* 27–44.

Brandon, Rajna Gibson, and Songtao Wang, 2013, Liquidity risk, return predictability, and hedge funds' performance: An empirical study, *Journal of Financial and Quantitative Analysis* 48, 219–244.

Breusch, T. S., 1978, Testing for autocorrelation in dynamic linear models, *Australian Economic Papers* 17, 334–355.

Brown, Stephen J, and William N Goetzmann, 2001, Hedge funds with style, Technical report, National Bureau of Economic Research.

Brown, Stephen J, and William N Goetzmann, 2003, Hedge funds with style, *The Journal of Portfolio Management* 29, 101–112.

Burghardt, Galen, and Lianyan Liu, 2013, Autocorrelation effects on cta and equity risk measurement, *The Journal of Alternative Investments* 16, 19–42.

Capocci, Daniel, and Georges Hübner, 2004, Analysis of hedge fund performance, *Journal of Empirical Finance* 11, 55 – 89.

Chung, Sam Y, Mark Rosenberg, and James F Tomeo, 2004, Hedge fund of fund allocations using a convergent and divergent strategy approach, *The Journal of Alternative Investments* 7, 44–53.

Cochrane, John H, 1988, How big is the random walk in gnp?, *The Journal of Political Economy* 893–920.

De Souza, Clifford, and Suleyman Gokcan, 2004, Hedge fund investing: A quantitative approach to hedge fund manager selection and de-selection, *The Journal of Wealth Management* 6, 52–73.

Edelman, Daniel, William Fung, and David A. Hsieh, 2013, Exploring uncharted territories of the hedge fund industry: Empirical characteristics of mega hedge fund firms, *Journal of Financial Economics* 109, 734 – 758.

Eling, Martin, 2009, Does hedge fund performance persist? overview and new empirical evidence, *European Financial Management* 15, 362–401.

Fung, William, and David A Hsieh, 1997, Survivorship bias and investment style in the returns of ctas, *The Journal of Portfolio Management* 24, 30–41.

Fung, William, and David A Hsieh, 2001, The risk in hedge fund strategies: theory and evidence from trend followers, *Review of Financial Studies* 14, 313–341.

Fung, William, and David A Hsieh, 2004, Hedge fund benchmarks: A risk-based approach, *Financial Analysts Journal* 65–80.

Getmansky, Mila, Andrew W. Lo, and Igor Makarov, 2004, An econometric model of serial correlation and illiquidity in hedge fund returns, *Journal of Financial Economics* 74, 529–609.

Godfrey, Leslie G, 1978, Testing against General Autoregressive and Moving Average Error Models When the Regressors Include Lagged Dependent Variables, *Econometrica* 46, 1293–1301.

Gregoriou, Greg N, Georges Hübner, and Maher Kooli, 2010, Performance and persistence of commodity trading advisors: Further evidence, *Journal of Futures Markets* 30, 725–752.

Greyserman, Alex, and Kathryn Kaminski, 2014, *Trend Following with Managed Futures: The Search for Crisis Alpha* (Wiley Trading, New York).

Hansen, Lars Peter, 1982, Large sample properties of generalized method of moments estimators, *Econometrica: Journal of the Econometric Society* 1029–1054.

Harlow, W. V., and Ramesh K. S. Rao, 1989, Asset pricing in a generalized mean-lower partial moment framework: Theory and evidence, *Journal of Financial and Quantitative Analysis* 24, 285–311.

Joenväärä, Juha, Robert Kosowski, and Pekka Tolonen, 2012, Revisiting 'stylized facts' about hedge funds-insights from a novel aggregation of the main hedge fund databases, *Unpublished working paper, Centre for Hedge Fund Research, Risk Management Laboratory, Imperial Colledge Business School* .

Jorion, Philippe, and Christopher Schwarz, 2013, The delisting bias in hedge fund databases, *The Journal of Alternative Investments* 16, 37–47.

Kaminski, Kathryn M, and Alexander Mende, 2011, Crisis alpha and risk in alternative investment strategies, Technical report, CME Group.

Kendall, M, and A Stuart, 1976, The advanced theory of statistics griffin: London 1979, *Vol. I* 102.

Khandani, Amir E., and Andrew W. Lo, 2011, Illiquidity premia in asset returns: An empirical analysis of hedge funds, mutual funds, and us equity portfolios, *Quarterly Journal of Finance* 01, 205–264.

Kosowski, Robert, Narayan Y. Naik, and Melvyn Teo, 2007, Do hedge funds deliver alpha? a bayesian and bootstrap analysis, *Journal of Financial Economics* 84, 229–264.

Ledoit, Oliver, and Michael Wolf, 2008, Robust performance hypothesis testing with the sharpe ratio, *Journal of Empirical Finance* 15, 850 – 859.

Levich, Richard M, and Rosario C Rizzo, 1999, Alternative tests for time series dependence based on autocorrelation coefficients, *Unpublished Working Paper, New York University Salomon Center S* .

Lo, Andrew W, 2002, The statistics of sharpe ratios, *Financial Analysts Journal* 36–52.

Lo, AW, and AC MacKinlay, 1988, Stock market prices do not follow random walks: evidence from a simple specification test, *Review of Financial Studies* 1, 41–66.

Malkiel, Burton G, and Atanu Saha, 2005, Hedge funds: risk and return, *Financial analysts journal* 61, 80–88.

Moskowitz, Tobias J., Yao Hua Ooi, and Lasse Heje Pedersen, 2012, Time series momentum, *Journal of Financial Economics* 104, 228 – 250, Special Issue on Investor Sentiment.

Rzepczynski, Mark S, 1999, Market vision and investment styles: Convergent versus divergent trading, *The Journal of Alternative Investments* 2, 77–82.

Sadka, Ronnie, 2010, Liquidity risk and the cross-section of hedge-fund returns, *Journal of Financial Economics* 98, 54–71.

Schneeweis, Thomas, Richard Spurgin, and David McCarthy, 1997, Informational content in historical cta performance, *Journal of Futures Markets* 17, 317–339.

Sortino, Frank A, and Robert Van Der Meer, 1991, Downside risk, *The Journal of Portfolio Management* 17, 27–31.

Sun, Zheng, Ashley Wang, and Lu Zheng, 2012, The road less traveled: Strategy distinctiveness and hedge fund performance, *Review of Financial Studies* 25, 96–143.

Titman, Sheridan, and Cristian Tiu, 2011, Do the best hedge funds hedge?, *Review of Financial Studies* 24, 123–168.

White, Halbert, 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica: Journal of the Econometric Society* 817–838.

Chapter 3

Intraday momentum in FX markets: disentangling informed trading from liquidity provision[1]

Journal of Financial Markets
In Press

3.1 Introduction

Market participants need time to interpret and react to new information. Consequently, the dissemination of news potentially leaves room for predictability over short horizons of time. Theoretically, participants' trades are likely to be informative of future returns, given that they contain private information (Lyons, 1995).

A number of papers show that interdealer order flow in foreign exchange (FX) markets is indeed predictive of future returns. Payne (2003) shows that trades carry information and have a substantial permanent impact on prices. Similarly, Chordia et al. (2005) show that order flow

[1]This chapter is based on joint work with Kevin Lampaert (Ghent University) and Michael Frömmel (Ghent University).

is predictive of future returns over the very short horizon. More recently, Chordia et al. (2008) find that very short-term predictability is diminished when bid-ask spreads are narrower, indicating that liquidity enhances market efficiency through increased arbitrage activity. This finding suggests that liquidity also plays a role in the short-term predictability of returns.

Although most of the above studies focus on very short horizons, Gao et al. (2015) take a considerably longer perspective while staying in the field of intraday high-frequency data. In particular, they investigate the predictability of a security's first half-hour return on its last half-hour return and find that the former is positively predictive of the latter. This finding suggests that, in addition to predictability over very short periods of time, there also appears to be predictability over considerably longer periods of time during the trading day. To date, however, no researchers have empirically tested the likely drivers of this "intraday momentum".

Our contribution to the literature on FX microstructure is twofold. First, by using a long sample of transaction-level FX market data at tick frequency, we construct high-frequency measures of the likely drivers of intraday momentum in the ruble market. Using these measures, we analyze whether intraday momentum is stronger on days with more informed trading or when demand for liquidity is higher. These hypotheses capture the likely explanations of how market participants' behavior may generate the observed intraday momentum effect.

For the RUB-USD FX market, and contrary to the results of Gao et al. (2015) for the equity market, we do not find any evidence supporting the idea that intraday momentum is the result of strategic informed trading during the opening and closing of the trading session. This finding is consistent with the earlier finding that informed traders in the RUB-USD FX market mainly trade during the opening of the trading sessions in the Moscow Interbank Currency Exchange (MICEX) (Menkhoff and Schmeling, 2010). Instead, our results for the ruble market indicate that opening half-hour returns are positively predictive of closing half-hour returns on days when bid-ask spreads are high during the opening half-hour. We hypothesize that high spreads are consistent with higher levels of liquidity provision by some market participants following heavy trading early in the morning. Taken together, our results lend support to the argument that risk aversion to overnight holdings and a potential disposition effect among liquidity-providing market participants drive intraday momentum in the ruble market.

Second, our findings also contribute to a better understanding of intraday momentum along several other dimensions. In particular, we corroborate the finding of Gao et al. (2015) that the trading hours of the non-major currency's domestic market matter for intraday momentum. Although these authors observe a general lack of intraday momentum in major currencies vis-à-vis the U.S. dollar when considering U.S. trading hours, they find some weak evidence of intraday momentum when they determine implicit trading hours, based on increases in volume in international equity index futures. Our results for the RUB-USD currency pair show that, by considering the *explicit* trading hours of the MICEX, significant levels of intraday momentum are present. Clearly, the explicit nature of the trading hours helps to identify the relevant periods over which intraday momentum occurs in this FX market. Finally, our results also support the earlier observation that intraday momentum is more pronounced during financial crisis periods.

The remainder of this paper is structured as follows. In Section 3.2, we provide an overview of the related literature and formulate the different mechanisms that may drive intraday momentum. In Section 3.3, we describe the data used for our empirical analysis. In section 3.4, we outline the concept of intraday momentum and present the methodology used to measure the degree of informed trading and liquidity demand. In section 3.5, we discuss the results. In section 3.6, we assess the robustness of the results. We conclude in Section 3.7.

3.2 Motivation and related literature

Gao et al. (2015) suggest two potential mechanisms that may drive intraday momentum in financial markets. First, the intraday pattern can be the result of liquidity provision by some market participants (e.g., day traders, market makers, etc.). With price dissemination being the highest at the beginning of a trading session (Bloomfield et al., 2005) when market participants react to macroeconomic news released overnight before the start of the trading session, temporary imbalances may arise when market participants react similarly to news. Day traders and market makers may be motivated to take opposite positions to provide liquidity to the market. However, although these traders may quickly close out winning positions throughout the day, they may be more reluctant to rapidly close out losing positions. However, the prospect of having to hold positions overnight may convince traders and market makers to close out the positions nonetheless. Gao et al. (2015) point to a disposition effect among (day) traders (Odean, 1998; Locke and Mann, 2005) to motivate such asymmetric behavior. The risk management practices of financial institutions, however, may similarly force traders to close out positions before the end of the day. This behavior of (foreign exchange) dealers' offloading undesired inventory has been widely documented in the literature (Lyons, 1995; Bjønnes and Rime, 2005).

Second, intraday momentum is also theoretically consistent with the strategic behavior of informed traders. Theoretically, Kyle (1985) and Admati and Pfleiderer (1988) argue that informed traders will time their trades during high-volume periods to hide their informational advantage and to limit the price impact. Doing so will force informed traders to trade during high-volume periods (see Bloomfield et al., 2005). Given the well-known U-shape in intraday trading volume, the implication is that they will trade at the beginning and near the end of the trading day. If informed traders indeed place their trades during periods of heavy trading and if their trading has a (permanent) price impact, then this may also drive the observed predictability in intraday returns.

Both explanations are closely related to the existing FX microstructure literature on the predictability of returns in FX markets. Research indicates that fundamentals, proxied with macroeconomic variables, perform poorly in forecasting future exchange rate movements (e.g., Evans and Lyons, 1999); however, this is not the case for order flow and liquidity. In particular, it is well founded that order flow is predictive of returns over the very short term. For example, Payne (2003) shows that market participants' trades carry information and have a substantial permanent impact on prices. Similarly, Chordia et al. (2005) show that order flow is predictive of future returns over the very short horizon.

Theoretically, the predictability of future returns based on order flow is consistent with strategic order splitting among informed traders. Given that information among market participants is heterogeneous, some participants are likely to participate in strategic trading to disguise their superior information. One way to lower the impact of their trades is through order splitting (Chakravarty, 2001), which results in correlated trades.

Love and Payne (2008) show that there is short-term predictability through order flow when public information is released, which suggests that the predictability is driven by information processing. Simultaneously, Evans and Lyons (2005) show that FX markets incorporate news only gradually, over the matter of a few days, rather than instantaneously. Similarly, Rime et al. (2010) confirm gradual learning and show that order flow is a strong predictor for daily returns. The above literature indicates that both transitory and permanent price impacts seem to be predictable from past order flow, at least over short horizons.

There are recent reports that liquidity is also an important explanatory variable in the price discovery process. Chordia et al. (2008) find that very short-term predictability is diminished when bid-ask spreads are narrower, indicating that liquidity enhances market efficiency through increased arbitrage activity. More recently, Boudt and Petitjean (2014) show that changes in order imbalances are informative of price discovery. This finding suggests that liquidity also plays a role in the short-term predictability of returns.

3.3 Data description and institutional features

3.3.1 Data

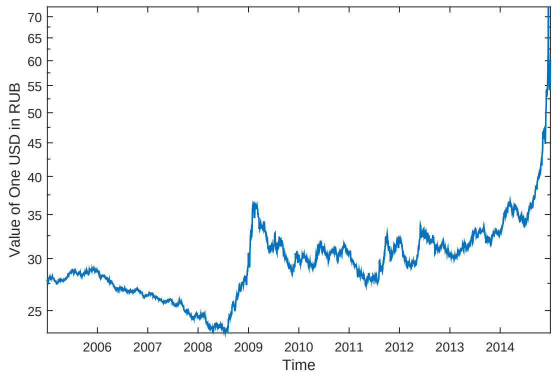
We use a particularly long-time span of intraday transaction-level data at tick frequency on the Russian ruble-United States dollar. We obtain the data from the MICEX, the largest currency exchange in Russia and Eastern Europe. Spot trading in the RUB-USD currency pair equals 1.66% of total FX spot trading volume in 2013, meaning that the currency pair ranks as the 12th mostly heavily traded globally.

We obtain data for the January 12, 2005 to December 30, 2014 period. Although constrained to one particular currency pair, the data set offers several advantages. First, a long data span avoids a number of short sample problems that researchers often encounter in the microstructure literature, such as possible structural breaks or biases in the estimated parameters. Second, the sample period features both the 2007-2009 Global Financial Crisis and the more recent 2014 Russian currency crisis, during which the ruble was the object of the crisis. Figure 3.1 illustrates the evolution of the RUB-USD exchange rate over the sample period.

Both the 2007-2009 Global Financial Crisis and the 2014 Russian currency crisis are clearly discernible in Figure 1, with both instances leading to a meaningful depreciation in the value of the ruble versus the dollar. The figure also suggests somewhat higher volatility post-2008 compared to the first couple of years of the sample period.

The MICEX trading platform was jointly developed with Reuters and has features similar to the platform of Reuters or Electronic Brokerage Services (EBS). Participants can observe the price, the trading volume,

Figure 3.1: Evolution U.S. dollar - Russian ruble (2005-2014)



and the bid and ask prices with standing volumes. In contrast to most other FX markets, it is only possible to submit limit orders to the platform. However, market orders can be synthetically created by submitting marketable limit orders. The MICEX covers all domestic spot trading in Russia. Offshore trading in the RUB-USD is performed through and limited to non-deliverable forward contracts. To illustrate the fact that both platforms are very similar and that the MICEX is the main exchange for spot trading in RUB-USD worldwide, we note that trading on Thomson Reuters is transmitted to the MICEX during trading hours when the MICEX is open. Refer to Menkhoff and Schmeling (2010) for further details on ruble trading on the MICEX.

The data set contains the following information for every trade executed on the MICEX; a time-of-day time stamp (to the millisecond), the price at which the order is executed, and the size of the trade. Simultaneously, we also have information on the best bid- and ask price at the time every order is executed. From the transaction-level data, we calculate half-hour (30 minutes) log returns for each trading day t as follows:

$$r_{j,t} = log\left(\frac{p_{j,t}}{p_{j-1,t}}\right), \tag{3.1}$$

where $r_{j,t}$ represents the half-hour return at day t for intraday interval j and $p_{j,t}$ represents the exchange rate at day t (the value of one dollar quoted in rubles) at the end of intraday interval j. The first half-hour

Table 3.1: Summary statistics RUB-USD exchange rate

	Panel A: Full Sample Period (2005-2014)		Panel B: Financial Crises (2007-2009 & 2014)	
	First Half-hour Returns	Last Half-hour Returns	First Half-hour Returns	Last Half-hour Returns
Mean	-0.001%	0.004%	0.004%	0.004%
St. Dev.	0.589%	0.124%	0.798%	0.159%
Skewness	-1.842	-5.204	-2.278	-7.060
Kurtosis	56.896	138.337	45.945	132.897
Min	-8.932%	-2.943%	-8.932%	-2.943%
Max	6.265%	1.218%	6.265%	0.735%
# of Obs.	2,342	2,342	922	922

This table reports summary statistics for the RUB-USD exchange rate. We report statistics for both the first and the last half-hour return. Panel A contains the statistics for the full sample period 2005-2014, while Panel B contains the statistics for the crisis periods (2007-2009 & 2014).

return of each day is calculated based on the previous day's closing price. This way we also capture the overnight return component, which might drive the informed trading and liquidity demand we wish to analyze. At the same time, by using the previous day's closing price we avoid relying on the opening price. This is an important consideration, since the opening price is prone to pricing errors that may bias opening returns (see Amihud and Mendelson, 1987). Table 3.1 reports the summary statistics for the first and last half-hour returns we use. We report statistics both for the full sample period and for the crisis periods separately.

We observe that opening half-hour returns are considerably more variable than closing half-hour returns, which reflects information processing at the start of the trading session. In addition, both return series are negatively skewed, suggesting that large negative returns are considerably more prevalent than large positive returns.

3.3.2 Institutional features

The data set we consider has several features. First, and specific to the MICEX, the exchange changed the opening and closing hour on several occasions over the sample period. In all instances, the change in trading hours led to an increase in the number of hours that the MICEX is open.

Table 3.2: Overview trading sessions on the MICEX exchange for the RUB-USD

Period	Opening	Closing
01/01/2005 - 11/11/2008	10:00	14:00
12/11/2008 - 12/04/2013	10:00	15:00
13/04/2013 - 31/12/2014	10:00	17:00

Trading hours in Moscow local time (GMT+3).

Table 3.2 provides an overview of the changes in trading hours.

The changes in the number of trading hours imply that the amount of time between the first half-hour return and the last half-hour return, the returns of interest, is not constant throughout the sample period. Because intraday momentum is expected to occur mainly during the start and the end of the trading day, however, we expect that the phenomenon is unaffected by the particular time of day with which the trading half-hours correspond.

Second, we note that foreign exchange markets are generally considered to be open virtually around the clock, with at least one major exchange trading the major currency pairs virtually at any point in time during the week. As such, the notion of first half-hour and last half-hour returns in the case of foreign exchange markets may seem inappropriate. Although this is true, trading intensifies considerably when a currency's domestic financial market commences trading. Furthermore, returns, spreads, and volatility are impacted by the market activity of various financial centers (Andersen and Bollerslev, 1997). Therefore, it can be argued that foreign exchange markets generally have implicit opening and closing trading hours. In the case of our data set, trading in the currency pair is organized during a fixed trading session, providing us with explicit opening and closing hours.

Nonetheless, to the extent that market participants trade outside the trading hours of the MICEX, this particular feature of the FX market may work against finding intraday momentum. Simultaneously, both explanations for intraday momentum crucially depend on liquidity considerations. Thus, if the observed intraday momentum described above is driven by the particular behavior of traders suggested by both explanations, then they will likely trade during the trading hours of the MICEX.

Finally, we also briefly consider the particular institutional circumstances implied by FX markets. It is well known that trading on these markets is reserved to major banks and large institutions. This direct trading between major dealers covers the vast majority of foreign exchange traded volume and is often referred to as the first tier or wholesale tier. Our data set covers the trades executed on this wholesale tier market. Retail investors, mutual funds, and large non-financial firms are, however, not directly active on this tier. Instead, these investors transact bilaterally with banks or brokers who provide quotes. Depending on the inventories of the banks and brokers with which these investors transact, these investors' orders may or may not be passed on to the wholesale tier. This particular market structure means that retail investors, mutual funds, and large non-financial firms will only indirectly impact the foreign exchange market. As such, it is ultimately the manner in which market makers pass the resulting inventory changes to the wholesale tier that matters. We suggest that, if the liquidity needs of investors in the retail tier are large enough to materially impact the inventories of the market makers, then the effect will propagate to the trading on the wholesale tier. Despite the trading that follows from the two-tier structure of foreign exchange markets, trading on the wholesale tier strongly outweighs trading on the retail tier. The forces driving intraday momentum can be at play between participants in the wholesale tier, and we directly observe (the price impact of) this trading in our sample.

We conclude that the particular structure of FX markets does not, a priori, rule out the possibility of intraday momentum in foreign exchange markets, although some features likely work against observing an intraday momentum effect.

3.4 Methodology

To determine the existence of intraday momentum, we closely follow the approach used by Gao et al. (2015) and estimate predictive regressions. These authors note that the predictive regressions correspond to autoregressive (AR) models. Although this is true, changes to the trading hours by the MICEX over the sample period imply that, in our case, the exact lag length of the AR model varies over time (see Section 3.3). We therefore express the predictive regression as follows:

$$r_{l,t} = \alpha + \beta r_{f,t} + \epsilon_t, \tag{3.2}$$

where $r_{f,t}$ is the first half-hour return, $r_{l,t}$ is the last half-hour return and ϵ_t is the error term. We also consider the predictive value of the penultimate return, which we denote as $r_{sl,t}$. The inclusion of this term allows us to control for any short-term persistence in the exchange rate during the day and to isolate the predictive value of the last half-hour return.

To investigate the relation between informed trading and intraday momentum, we construct the dynamic probability of informed trading (DPIN) measure suggested by Chang et al. (2014). This measure builds on the empirical work of Campbell et al. (1992) and Avramov et al. (2006) and allows us to measure the degree of informed versus uniformed trading based on high-frequency transaction-level data. More specifically, this approach allows us to measure and track the presence of informed trades throughout the trading day based on a high frequency. The fact that financial markets are becoming increasingly computer-driven – potentially making private information increasingly short-lived – makes measuring informed trading at the intraday level increasingly important. The approach of Chang et al. (2014) allows us to avoid a degradation to lower frequencies of the PIN measure originally proposed by Easley et al. (1997).

Following Chang et al. (2014), we first perform a regression to isolate the unexpected half-hour return component (ϵ_t) from the return series while controlling for day-of-the-week effects (using dummy variables denoted D_j^{day}), time-of-day-effects (using dummy variables denoted D_j^{int}), and lagged half-hour returns (r_{t-k})[2]:

$$r_t = \alpha_0 + \sum_{i=1}^{4} \alpha_{1i} \cdot D_i^{day} + \sum_{j=1}^{J} \alpha_{2j} \cdot D_j^{int} + \sum_{k=1}^{12} \alpha_{3k} \cdot r_{t-k} + \epsilon_t. \tag{3.3}$$

Autocorrelation patterns in unexpected returns (or a lack thereof) indicate the presence of uninformed (informed) trading. In particular, Avramov et al. (2006) note that trades that take liquidity generate (future) price reversals. At the same time, sell trades in the presence of positive unexpected returns do not exhibit any autocorrelation and therefore indicate informed trading. Chang et al. (2014) argue that this can be extended to buy-side trades. The authors point out that buy-side trades in the presence of negative unexpected returns do not exhibit any autocorrelation, which again implies informed trading. Following Chang et al. (2014) our measure of informed trading is calculated as follows:

[2]Where J equals the number of intraday half-hour intervals in the specific period.

$$DPIN_t = \frac{NB_t}{NT_t} \cdot (\epsilon_t < 0) + \frac{NS_t}{NT_t} \cdot (\epsilon_t > 0), \tag{3.4}$$

where NB_t, NS_t, and NT_t are the number of buy, sell, and total trades, respectively, made during the half-hour interval from t to $t-1$ and $(\epsilon_t < 0)$ and $(\epsilon_t > 0)$ are sign indicators that equal one when the unexpected return is smaller and larger than zero, respectively, and zero otherwise.

To analyze the alternative explanation, i.e., whether liquidity provision to some extent drives intraday momentum, we require a measure that identifies the trading days in which market participants can be expected to provide liquidity to the market. For purposes of analysis, we focus on the tightness dimension of liquidity (Kyle, 1985). This is the main dimension of liquidity and is measured using the equal-weighted quoted spread (EWQS). This metric measures the average bid-ask spread over a given period of time. We hypothesize that, on days where the EWQS was higher during the first half-hour, more liquidity was demanded by market participants (e.g., as a consequence of economic news that was released overnight), meaning that some day traders or market makers are more likely to have provided the required liquidity.

3.5 Results

In this section, we first establish the presence of intraday momentum and assess the economic significance of the effect. Then we explore the relation between intraday momentum, informed trading, and liquidity demand.

3.5.1 Intraday momentum in RUB-USD

We start by running a set of predictive regressions in the spirit of Gao et al. (2015). In particular, we explore whether the first half-hour return, the penultimate half-hour return, and a combination of both independent variables are predictive of the last half-hour return. The results are reported in Table 3.3.

The results for the entire sample, reported in Panel A of Table 3.3, indicate that there is no significant relation between the last half-hour return and the first half-hour return. Although the coefficient has the expected sign, it is not significant at conventional levels, with a p-value of 0.12. The results for the penultimate half-hour return are similar, although the relation appears to be even weaker. When we include both

Table 3.3: Predictability of last half-hour return

	Panel A: Full Sample			**Panel B: Crises (2007-2009 & 2014)**			**Panel C: Excluding Crises**		
Variables	r_l	r_l	r_l	r_l	r_l	r_l	r_l	r_l	r_l
r_f	0.0428		0.0412*	0.0698*		0.0656**	-0.0097		-0.0097
	(0.028)		(0.025)	(0.038)		(0.031)	(0.011)		(0.011)
r_{sl}		-0.1642	-0.1493		-0.2716	-0.2271		0.0020	0.0033
		(0.148)	(0.124)		(0.234)	(0.178)		(0.054)	(0.053)
Intercept	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	2,342	2,342	2,342	922	922	922	1,420	1,420	1,420
R^2 (%)	4.3	1.9	5.9	12.2	5.1	15.7	0.2	0.0	0.2

This table reports the results for the sample period from January 12, 2005 to December 30, 2014 by regressing the closing half-hour return (r_l) on the first half-hour return (r_f) and the second last half-hour return (r_{sl}). Panel A contains the results for the full sample period, whereas Panel B reports the results for the crisis periods. Panel C contains the results for the non-crisis periods. Newey and West (1987) robust standard errors in parentheses. Significance at the 1%, 5%, and 10% levels indicated by ***, **, and *, respectively.

intraday returns in the predictive regression, however, the coefficient on the first half-hour return becomes significant at conventional levels, albeit only at the 10% level. One potential reason could be microstructural issues such as bid-ask bounces, which cause intraday returns to exhibit mean-reverting behavior over short intervals. These results, although suggestive, are somewhat thin.

Second, we examine whether the relation differs during periods of financial stress. We classify the 2007-2009 Global Financial Crisis and the 2014 Russian currency crisis as periods of financial distress. The results, reported in Panels B and C of Table 3.3, indicate that intraday momentum is considerably more pronounced during periods of financial stress. During non-crisis periods, however, the relation does not appear significant. This finding is consistent with the findings of Gao et al. (2015), who find that intraday momentum is more pronounced during the 2007-2009 Global Financial Crisis.

Third, to test the predictive ability of intraday momentum out-of-sample (OOS), we also perform OOS forecasts. In particular, we run the above predictive regression with expanding windows, adding one day at a time. Using the estimated coefficients of the predictive regression (denoted using hats) and the value of the predictive variable at time s, we can generate a forecast of the return at time $s+1$:

$$\hat{r}_{l,s+1} = \hat{\alpha} + \hat{\beta} r_{f,s}. \tag{3.5}$$

We perform these estimations for $s = s_0, ..., t-1$, thus generating a time series of OOS return forecasts. s_0 is the initial sample size used to estimate the model (in our application, four years). We then estimate the OOS $\mathbf{R^2}$ to measure OOS forecastability:

$$OOS\, R^2 = 1 - \frac{\frac{1}{T-s_0}\sum_{s=s_0}^{T-1}\left(r_{l,s+1} - \hat{r}_{l,s}\right)^2}{\frac{1}{T-s_0}\sum_{s=s_0}^{T-1}\left(r_{l,s+1} - \bar{r}_{l,s}\right)^2}, \tag{3.6}$$

where $\bar{r}_{l,s}$ is the historical mean of the last half-hour return, calculated from the expanding window of last half-hour returns. To test the significance of the OOS $\mathbf{R^2}$, we employ the F-statistic of McCracken (2007). In Table 3.4, we report the results for the OOS $\mathbf{R^2}$.

Similarly to Gao et al. (2015), we obtain a significant OOS $\mathbf{R^2}$ of approximately 1.6%. This level of OOS $\mathbf{R^2}$ is very substantial compared to other works (e.g., Campbell and Thompson, 2008; Ferreira and Santa-Clara, 2011). Simultaneously, the penultimate return does not seem to have any OOS predictive power.

Table 3.4: OOS predictability

	OOS R^2	MSE-F
r_f	1.609%	21.948***
r_{sl}	-0.086%	-1.151
r_f and r_{sl}	1.640%	22.371***

This table reports the out-of-sample predictability results of the last half-hour by the first half hour return and the second-to-last half-hour return, using a set of recursive regressions. The initial sample period (s_0) is four years (2005-2008). Asterisks indicate statistical significance of the OOS R^2 using the MSE-F test

$$MSE - F = (T - s_0)\left(\frac{MSE_m - MSE_p}{MSE_p}\right).$$

Asymptotic critical values for the MSE test provided by McCracken (2007) used to test significance. Significance at the 10%, 5%, and 1% levels given by *, **, and ***, respectively.

A second method of testing the economic significance of the results is by analyzing the returns accruing to a simple market timing strategy that uses signals based on the first half-hour return. In particular, every trading day we take a long or short position at the beginning of the final half-hour period, depending on the return of the opening half-hour, and close out the position at the end of the trading day. We benchmark the performance of this particular strategy to a constant long strategy that always goes long at the beginning of every final half-hour and that closes out the position at the end of every trading day.[3]

The results in Table 3.5 indicate that, at least for the full sample period, the market timing strategy does not outperform the always long strategy. Interestingly, however, the returns to the intraday momentum strategy are positively skewed. This finding is in contrast to the always long series which, similar to the original first and last half-hour returns, is strongly negatively skewed. The disappointing performance of the strategy over the full sample matches the earlier observation that intraday momentum appears to be more pronounced during financial crises.

When we restrict the sample to the two crisis periods defined above, the market timing strategy performs particularly well. The strategy posts a higher return, a higher Sharpe ratio, and a higher success rate than the always long strategy. Interestingly, the returns to the intraday momentum

[3]We note that the returns to both strategies are comparable because both strategies have identical turnover and thus incur similar levels of transactions costs.

Table 3.5: Performance intraday momentum market timing strategy

	Panel A: Intraday Momentum Strategy		**Panel B: Always Long Strategy**	
	Full Sample	Crises	Full sample	Crises
Mean return	0.001%	0.009%	0.004%	0.004%
Sharpe	0.426	2.637	1.261	1.124
Skewness	5.196	7.279	-5.413	-7.060
Kurtosis	137.582	131.327	138.342	132.897
Success rate	49.530%	51.410%	52.135%	51.193%

This table reports summary statistics on the performance of a market timing strategy based on intraday momentum and an always-long trading strategy. The market timing strategy goes long when the first half-hour return is positive, and short otherwise. The always-long strategy always goes long the last half-hour of the trading day. The results are reported for the full sample and for the crisis periods.

strategy are again positively skewed, whereas the always long strategy exhibits negative skewness. As such, the intraday momentum trading strategy appears to limit downside risk.

Overall, these findings suggest that, although this fairly naïve market timing strategy does not generate attractive returns overall, the market timing strategy does appear to generate attractive returns in bad market states.

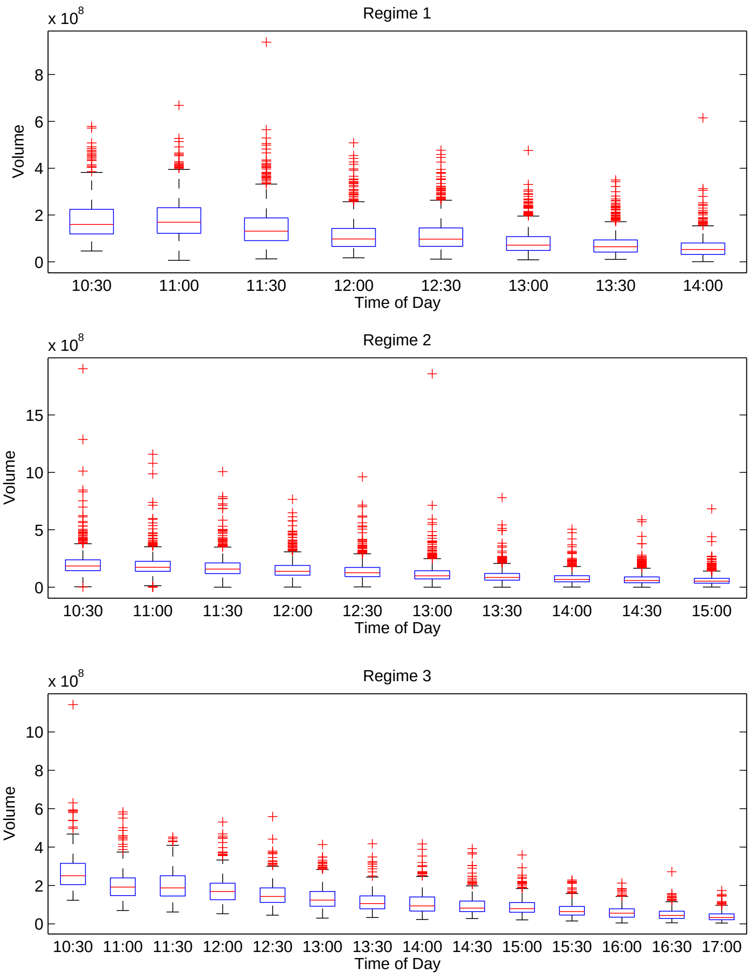
3.5.2 Informed trading versus liquidity provision

Having established the presence of intraday momentum in the RUB-USD market, we explore the likely drivers of intraday momentum outlined in the introduction. We first analyze how volume is distributed over the trading day. In Figure 3.2 we report the average half-hour trading volume (in USD) for the different trading hour regimes.[4]

Figure 3.2 shows that volume, on average, does not exhibit a U-shape, as is typical in equity markets (e.g., Jain and Joh, 1988). The box plots indicate that there is nevertheless considerable time series variation in the volume traded during every half-hour of trading. The fact that the

[4]For completeness, we report similar figures for DPIN and EWQS in the Appendix.

Figure 3.2: Distribution of volume (in U.S. dollars) over the trading day



RUB-USD market does not exhibit a U-shaped distribution in volume over the trading day has an important implication for the "informed trading hypothesis". This suggests that, although we find intraday momentum, informed trading may not be the main driver because there is generally no reason for informed traders to postpone their trading to the last half-hour of the trading day. This idea is consistent with the finding of Menkhoff and Schmeling (2010), who, using a short sample of data on the MICEX that includes anonymized trader identifiers, find that informed traders mainly trade during the opening of the trading sessions in the MICEX. Naturally, informed traders may have other considerations in addition to the trading volume for spreading trades over the trading day.

To formally analyze the relation between intraday momentum, informed trading, and liquidity demand, we estimate several model specifications. To be concise, we focus on the two crisis periods, for which we find intraday momentum to be most pronounced.[5] For purposes of comparison, we first repeat the baseline predictive regression of interest. The results are reported in column (1) of Table 3.6.

In Table A.2 of the Appendix, we observe that intraday momentum is related to the realized volatility and trading volume over the first half-hour of trading.[6] To control for both effects, we include the realized volatility during the first half-hour and the (common log of) volume as controls in the regression and report the results in column (2). Controlling for volume and realized volatility, we observe no change in the sign, magnitude, or significance of the estimated coefficients. For completeness, we report the pairwise correlations between the variables of interest in Table A.3. of the Appendix.[7]

Turning to the other specifications, column (3) of Table 3.6 reports the results for the specification examining the relation between intraday momentum and periods of low and high levels of informed trading. In particular, we construct a set of dummy variables that equal 1 depending on whether the level of informed trading during the first half-hour is in the top (D_H), middle, or bottom (D_L) tercile, respectively. We then

[5]The results for the full sample, reported in Table A.1 of the Appendix, remain qualitatively the same.

[6]Gao et al. (2015) show that intraday momentum is positively associated with volume and volatility. We repeat their analysis and find that intraday momentum is positively associated with volume and volatility (see Table A.2 of the Appendix).

[7]The pairwise correlation between the EWQS and DPIN is high (0.69). However, the coefficients for the specifications in which we omit one of the two variables (cfr. infra) do not change meaningfully (see Table A.3), suggesting that multicollinearity is not an issue.

Table 3.6: Disentangling liquidity and informed trading during crises

Variables	(1) r_l	(2) r_l	(3) r_l	(4) r_l	(5) r_l
r_f	0.0656**	0.0608**	0.0954*	0.0071	0.0368
	(0.031)	(0.025)	(0.051)	(0.017)	(0.037)
r_{sl}	-0.2271	-0.2467	-0.2299	-0.2458	-0.2338
	(0.178)	(0.168)	(0.144)	(0.162)	(0.142)
$D_L(DPIN) \cdot r_f$			-0.0447		-0.0354
			(0.059)		(0.056)
$D_H(DPIN) \cdot r_f$			-0.0756		-0.0765
			(0.056)		(0.054)
$D_L(EWQS) \cdot r_f$				0.0136	0.0214
				(0.027)	(0.027)
$D_H(EWQS) \cdot r_f$				0.0642*	0.0671**
				(0.036)	(0.031)
Opening σ^2_{RV}		-0.0955	-0.0941	-0.0925	-0.0902
		(0.078)	(0.070)	(0.075)	(0.067)
Opening $log(Volume)$		-0.0000	-0.0000	-0.0000	-0.0000
		(0.000)	(0.000)	(0.000)	(0.000)
Intercept	0.0000	0.0002	0.0002	0.0002	0.0002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	922	922	922	922	922
R^2 (%)	15.7	19.0	21.7	20.2	23.0

This table presents regression results for the sub-sample that covers the 2007-2009 Global Financial Crisis and the 2014 Russian currency crisis. In the regression for the results in column (1), we regress the closing half-hour return (r_l) on the first half-hour return (r_f) and the second last half-hour return (r_{sl}). In the regression for the results in column (2), we control for volume and realized volatility during the first half-hour of trading. Column (3) reports the results for an evaluation of the impact of informed trading on the closing half-hour return. In column (4), we measure the impact of liquidity on the closing half-hour return. Finally, in the regression for the results in column (5) we combine both specifications. Newey and West (1987) robust standard errors in parentheses. Significance at the 1%, 5%, and 10% levels indicated by ***, **, and *, respectively.

interact these dummy variables with the observed return during the first half-hour of trading, omitting the middle tercile to serve as the baseline. The results in column (3) suggest that the predictive relation is not significantly stronger during periods of above-average or below-average levels of informed trading in the first half-hour of the trading day.

We also analyze the alternative hypothesis, which relates intraday momentum to liquidity provision by day traders during the start of the trading session. Similar to the case of informed trading, we divide all trading days into three different terciles, depending on the value of the EWQS over the opening half-hour interval. We report the results in column (4) of Table 3.6.

All else being equal, higher quoted spreads can also be the result of high volatility. However, because we include the realized volatility over the first half-hour of trading as a control variable, the regression specification in column (4) of Table 3.6 should control for this effect and allow us to better isolate the impact of liquidity provision following strong liquidity demand. In this regression, we also interact the resulting dummy variables with the first half-hour return. Interestingly, we find that the first half-hour return in column (4) becomes insignificant. Instead, the interaction term that interacts the first half-hour return with the dummy in periods of high quoted spreads becomes positive and significantly so. This finding suggests that intraday momentum is the result of high liquidity demand by market participants during the opening combined with dealers' risk aversion to overnight inventory. Finally, we control for the level of informed trading; see column (5). Menkhoff and Schmeling (2010) find that informed traders in the MICEX tend to trade when spreads are higher, implying that we need to control for the level of informed trading.

Interestingly, controlling for informed trading in column (5) of Table 3.6, we find that the relation becomes even more pronounced from a statistical perspective. This result suggests that intraday momentum tends to occur during trading days when quoted spreads are high, even when controlling for the potential effect of informed trading on spreads. We interpret this finding as supportive of the hypothesis that intraday momentum is to a certain extent driven by a high liquidity demand during the morning, combined with a strong risk aversion to overnight holdings potentially driven by risk management policies, the disposition effect or habits among market makers.

Are there institutional circumstances that may inform why intraday momentum in the ruble market appears to be the result of liquidity provision, rather than informed trading? The main differences between foreign exchanges and other financial markets are the sheer size of FX markets and the fact that these markets are only accessible by major dealers. We suggest that, because the FX market is considerably larger in terms of notional value, informed trading is less likely to impact prices. Simultaneously, however, if a sufficiently large fraction of the market's participants reacts similarly to a news announcement, then liquidity demand can be expected to meaningfully impact prices (albeit temporarily).[8]

Second, the results suggest that the traders who provide liquidity to these early trades close their positions and thus take exactly the same direction as the information-driven trades at the start of the day. Because these traders mirror the information-based trades in the morning, what is their motivation and why do they not adjust their behavior?

We note that in the microstructure theory, the bid-ask spread consists of three components: an order processing component, an adverse selection component, and an inventory holding component (Huang and Stoll, 1997). Changes in the bid-ask spread, in this case, are likely to be driven by changes in the latter two components.[9]

One reason why the intraday pattern, if it is indeed driven by liquidity provision during the opening session, may continue to exist is the following. We can assume that, when market makers set their prices, they will most likely take into consideration the ease with which they will be able to eliminate the position. As such, a market maker will be willing to

[8]A second reason why liquidity may be the prime driver of intraday momentum is the following. Informed traders attempt to hide their informational advantage by splitting large orders (Chordia and Subrahmanyam, 2004) into several smaller, medium-sized transactions (Chakravarty, 2001). Thus, their trading will be geared towards avoiding a meaningful price impact. To the extent that traders are successful at hiding their informational advantage, we will not observe any intraday momentum. Moreover, although excess inventories require trading near the end of the trading day, the informed trading hypothesis provides no rationale for informed traders to always trade in both the morning and the afternoon. Because informed traders want to monetize their informational advantage as quickly as possible (Bloomfield et al., 2005), it is less likely that they will want to wait until the end of the trading day, especially, in markets as deep as the FX markets. Moreover, earlier work using the same data on the same market concludes that FX traders on the MICEX mainly trade during the opening session through medium-sized orders (Menkhoff and Schmeling, 2010).

[9]The order processing component refers to market makers' fixed costs. The adverse selection component compensates the market maker in cases when he or she is trading against a counterparty who may have superior information. For example, aggressive (market) orders may indicate that the counterparty has private information and thus may motivate the market maker to increase the spread. Finally, the inventory holding component refers to a premium that the market maker requires for providing liquidity during periods of unbalanced flows.

provide liquidity provided that the premium (i.e., the inventory holding component) received is higher than the likely cost of having to liquidate the position later that day. In other words, the profit from providing liquidity during the first half-hour should offset the expected loss from forced liquidation later that trading day. This may provide explanatory power for why the effect persists and why traders who generate the effect continue to survive.

3.6 Robustness checks

We now present the results of additional regressions to test the robustness of the intraday momentum effect on several dimensions. In particular, we analyze whether the effect is robust across different subsamples, different return sampling frequencies, alternative definitions of liquidity, and changes in the estimation method.

3.6.1 Subsample analysis

We repeat the analysis for both crisis periods separately. If intraday momentum in the RUB-USD market is indeed primarily a crisis-based phenomenon, we should observe a significant relation during both crisis periods. We report the results for the 2007-2009 Global Financial Crisis and the 2014 Russian currency crisis in Panels A and B of Table 3.7, respectively.

Although the relation is significant in both instances, the results in Table 3.7 show that intraday momentum is especially pronounced during the 2014 Russian currency crisis. This finding should not come as a surprise, given that the ruble was to a large extent the object of the crisis. This was not the case during the 2007-2009 Global Financial Crisis, where equity and credit markets played the leading part.

3.6.2 Choice of the return frequency

The use of half-hour returns strictly follows earlier work on intraday momentum in financial markets. However, this usage leaves unanswered the question of whether the peak of momentum predictability indeed is situated around this particular frequency. A natural question that arises is whether the observed intraday momentum is robust to the use of dif-

Table 3.7: Robustness check - 2007-2009 Global Financial Crisis & 2014 Russian currency crisis

	Panel A: 2007-2009 Global Financial Crisis			**Panel B: 2014 Russian currency crisis**		
Variables	r_l	r_l	r_l	r_l	r_l	r_l
r_f	0.0214*		0.0214*	0.0926*		0.0820**
	(0.012)		(0.012)	(0.051)		(0.039)
r_{sl}		0.0053	0.0045		-0.4832	-0.3836
		(0.066)	(0.066)		(0.376)	(0.271)
Intercept	0.0001**	0.0001**	0.0001**	-0.0001	-0.0001	-0.0002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	686	686	686	236	236	236
R^2 (%)	1.4	0.0	1.4	19.7	12.3	27.2

This table presents the results for the sample periods of January 10, 2007 to December 30, 2009 and January 10, 2014 to December 30, 2014 regressing the closing half-hour return (r_l) on the first half-hour return (r_f) and the second last half-hour return (r_{sl}). Newey and West (1987) robust standard errors in parentheses. Significance at the 1%, 5%, and 10% levels indicated by ***, **, and *, respectively.

Table 3.8: Robustness check - sensitivity of intraday momentum to the return frequency

r_f/r_l	60 Minutes	30 Minutes	15 Minutes
60 Minutes	0.0457**	0.0667*	0.0245*
30 Minutes	0.0513**	0.0698*	0.0269**
15 Minutes	0.0214	0.0330*	0.0273**

This table presents regression results for the return frequency sensitivity analysis. The coefficients for the specification under equation (2) for alternative opening and closing return frequencies are displayed. Significance using Newey and West (1987) standard errors at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

ferent frequencies.[10] To test whether intraday momentum is sensitive to the frequency and whether half-hour returns are the peak of the observed predictability, we re-run the regression in equation (1) for different combinations of return frequencies. In particular, we perform $K \times K$ regressions to analyze all potential combinations of the first and final 15-minute, half-hour, and one-hour returns. We report the coefficients of interest in Table 3.8.

In Table 3.8, we find that intraday momentum is robust to the frequency employed. In particular, the price action at the start of the trading day is predictive of the price evolution near the end of the trading day, and the relation is robust to the particular interval chosen. In economic terms, the effect is strongest for opening half-hour returns on closing half-hour returns.

Next, we analyze the robustness of the main results to a change in frequency. Because both proposed mechanisms that may drive intraday momentum can be expected to be at play especially during the very start and end of the trading session, we re-run the main analysis, calculating all variables of interest over the first 15 minutes of trading, and try to predict the return during last 15 minutes of the trading session. The first column of Table 3.9 reports the results. Our findings continue to hold, indicating that the mechanism that drives intraday momentum is at play at the very start of the trading session.

[10] We thank an anonymous referee for calling attention to this point.

Table 3.9: Robustness check - alternative definitions and estimation method

Variables	(1) r_l	(2) r_l	(3) r_l	(4) r_l
r_f	0.0031	0.0507	0.0418	0.0368
	(0.019)	(0.036)	(0.038)	(0.036)
r_{sl}	-0.0826	-0.2285	-0.2312	-0.2338*
	(0.088)	(0.143)	(0.142)	(0.141)
$D_L(DPIN) \cdot r_f$	0.0035	-0.0413	-0.0346	-0.0354
	(0.020)	(0.058)	(0.055)	(0.055)
$D_H(DPIN) \cdot r_f$	-0.0307	-0.0743	-0.0763	-0.0765
	(0.027)	(0.055)	(0.054)	(0.054)
$D_L(EWQS) \cdot r_f$	0.0193			0.0214
	(0.019)			(0.027)
$D_H(EWQS) \cdot r_f$	0.0398**			0.0671**
	(0.018)			(0.031)
$D_L(ES) \cdot r_f$		0.0165		
		(0.037)		
$D_H(ES) \cdot r_f$		0.0491*		
		(0.027)		
$D_L(VWQS) \cdot r_f$			0.0049	
			(0.027)	
$D_H(VWQS) \cdot r_f$			0.0623**	
			(0.030)	
Opening σ^2_{RV}	-0.0827*	-0.0919	-0.0902	-0.0902
	(0.047)	(0.069)	(0.067)	(0.067)
Opening $log(Volume)$	-0.0000	0.0000	0.0000	0.0000
	(0.000)	(0.000)	(0.000)	(0.000)
Intercept	0.0003	0.0002	0.0002	0.0000
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	922	922	922	922
R^2 (%)	11.3	22.2	23.0	

This table reports the results for the robustness checks. Column (1) reports the results of the main specification using an alternative return frequency of 15-minutes for the first- and last half-hour return. Column (2) presents the results using the effective spread as a measure of liquidity. Column (3) similarly presents the results using the volume-weighted quoted spread as a liquidity measure. Finally, column (4) reports the results obtained from estimation of the main specification using a two-step GMM.Newey and West (1987) robust standard errors in parentheses in column (1), (2), and (3). Significance at the 1%, 5%, and 10% levels indicated by ***, **, and *, respectively.

3.6.3 Alternative liquidity measures

Next, we assess the robustness of our main results to different measures of liquidity. To that end, we repeat the specifications for Table 3.6 using several alternative measures of liquidity that we can construct from our data. First, we employ effective spread (ES) as the liquidity metric. The result is shown in column (2) of Table 3.9 and confirms our baseline results and the results described above. In particular, we continue to find that liquidity appears to be the main driver of intraday momentum in the RUB-USD FX market.

Second, we replace the EWQS variable from our baseline analysis with the volume-weighted quoted spread (VWQS). This measure weights the bid-ask spreads by the volume of trades, and therefore, it takes into consideration the size of the trade matching the observed bid and ask prices. We report the results in column (3) of Table 3.9. Here too, we find that the intraday momentum effect is stronger when bid-ask spreads are high during the opening half-hour.

3.6.4 Estimation method

The estimations we have performed so far are based on OLS. Return series, however, tend to exhibit volatility clustering, which, from a statistical perspective, induces heteroscedasticity. In addition, high-frequency data often exhibit significant levels of negative autocorrelation over very short intervals (Roll, 1984) and positive autocorrelation over slightly longer intervals. Some of these patterns are the result of microstructure-related issues such as the bid-ask bounce, whereas others follow from the fact that information processing takes time (Chordia et al., 2005). Using Newey and West (1987) robust standard errors, we have so far accounted for such effects on the estimation results.

Nonetheless, because we do not know the full shape of the distribution of the data, we re-estimate the main results using generalized method of moments (GMM). Although the moments we impose are identical to the moments under OLS, a two-step GMM allows us to efficiently estimate the model when we face heteroscedasticity and autocorrelation of an unknown form. We report the result in the final column of Table 3.9. The results indicate that our findings are robust to the particular estimation method employed.

3.7 Conclusion

In this paper, we use a long sample of transaction-level data at tick frequency on the Russian ruble-U.S. dollar currency pair from the MICEX to investigate the likely drivers of intraday momentum in this FX market.

We contribute to the emerging literature of momentum at the intraday level in several ways. First, we find no evidence that intraday momentum in the ruble market is the result of market participants' strategic trading during high-volume periods. Two observations motivate this conjecture. First, there is no reason for informed traders in the ruble market to postpone trading until the last half-hour of trading, given that volume in the market does not exhibit a U-shape intraday pattern. This is consistent with work by Menkhoff and Schmeling (2010), who find that informed traders in this particular market mainly trade during the opening of the trading session. Second, we do not find a stronger intraday momentum pattern on days with more informed trading in the first half-hour of trading.

Instead, we find evidence that closing half-hour returns are positively related to opening half-hour returns on days when spreads in the ruble market are high during the opening half-hour. These high spreads are consistent with a strong liquidity demand by market participants in the first half-hour of trading. This finding lends support to the argument that dealers and other liquidity providers in the ruble market are trying to offload unwanted inventories (Lyons, 1995; Bjønnes and Rime, 2005) due to their risk aversion to overnight holdings. This interpretation is consistent with the empirical findings of Bjønnes et al. (2005), who show that non-financial customers are the main liquidity providers in the overnight foreign exchange market.

Second, we provide additional evidence that corroborates the finding of Gao et al. (2015) that explicit trading hours matter for intraday momentum. The particular nature of the RUB-USD FX market, a currency pair for which spot trading is only possible on the MICEX, provides a unique case where FX trading is subject to explicit trading hours. Finally, our results lend further support to the finding that intraday momentum is more pronounced during financial crises.

Bibliography

Admati, Anat R, and Paul Pfleiderer, 1988, A theory of intraday patterns: Volume and price variability, *Review of Financial studies* 1, 3–40.

Amihud, Yakov, and Haim Mendelson, 1987, Trading mechanisms and stock returns: An empirical investigation, *The Journal of Finance* 42, 533–553.

Andersen, Torben G, and Tim Bollerslev, 1997, Intraday periodicity and volatility persistence in financial markets, *Journal of Empirical Finance* 4, 115–158.

Avramov, Doron, Tarun Chordia, and Amit Goyal, 2006, The impact of trades on daily volatility, *Review of Financial Studies* 19, 1241–1277.

Bjønnes, Geir Høidal, and Dagfinn Rime, 2005, Dealer behavior and trading systems in foreign exchange markets, *Journal of Financial Economics* 75, 571–605.

Bjønnes, Geir Høidal, Dagfinn Rime, and Haakon O Aa Solheim, 2005, Liquidity provision in the overnight foreign exchange market, *Journal of International Money and Finance* 24, 175–196.

Bloomfield, Robert, Maureen O'Hara, and Gideon Saar, 2005, The "make or take" decision in an electronic market: Evidence on the evolution of liquidity, *Journal of Financial Economics* 75, 165–199.

Boudt, Kris, and Mikael Petitjean, 2014, Intraday liquidity dynamics and news releases around price jumps: Evidence from the djia stocks, *Journal of Financial Markets* 17, 121–149.

Campbell, John Y, Sanford J Grossman, and Jiang Wang, 1992, Trading volume and serial correlation in stock returns, Technical report, National Bureau of Economic Research.

Campbell, John Y., and Samuel B. Thompson, 2008, Predicting excess stock returns out of sample: Can anything beat the historical average?, *Review of Financial Studies* 21, 1509–1531.

Chakravarty, Sugato, 2001, Stealth-trading: Which traders' trades move stock prices?, *Journal of Financial Economics* 61, 289–307.

Chang, Sanders S., Lenisa V. Chang, and F. Albert Wang, 2014, A dynamic intraday measure of the probability of informed trading and firm-specific return variation, *Journal of Empirical Finance* 29, 80–94.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2005, Evidence on the speed of convergence to market efficiency, *Journal of Financial Economics* 76, 271–292.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2008, Liquidity and market efficiency, *Journal of Financial Economics* 87, 249–268.

Chordia, Tarun, and Avanidhar Subrahmanyam, 2004, Order imbalance and individual stock returns: Theory and evidence, *Journal of Financial Economics* 72, 485–518.

Easley, David, Nicholas M. Kiefer, and Maureen O'Hara, 1997, The information content of the trading process, *Journal of Empirical Finance* 4, 159–186.

Evans, Martin DD, and Richard K Lyons, 1999, Order flow and exchange rate dynamics, Technical report, National bureau of economic research.

Evans, Martin DD, and Richard K Lyons, 2005, Do currency markets absorb news quickly?, *Journal of International Money and Finance* 24, 197–217.

Ferreira, Miguel A., and Pedro Santa-Clara, 2011, Forecasting stock market returns: The sum of the parts is more than the whole, *Journal of Financial Economics* 100, 514–537.

Gao, Lei, Yufeng Han, Sophia Zhengzi Li, and Guofu Zhou, 2015, Intraday momentum: The first half-hour return predicts the last half-hour return, *Available at SSRN 2440866* .

Huang, Roger D, and Hans R Stoll, 1997, The components of the bid-ask spread: A general approach, *Review of Financial Studies* 10, 995–1034.

Jain, Prem C., and Gun-Ho Joh, 1988, The dependence between hourly prices and trading volume, *Journal of Financial and Quantitative Analysis* 23, 269–283.

Kyle, Albert S., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–1335.

Locke, Peter R., and Steven C. Mann, 2005, Professional trader discipline and trade disposition, *Journal of Financial Economics* 76, 401–444.

Love, Ryan, and Richard Payne, 2008, Macroeconomic news, order flows, and exchange rates, *Journal of Financial and Quantitative Analysis* 43, 467–488.

Lyons, Richard K, 1995, Tests of microstructural hypotheses in the foreign exchange market, *Journal of Financial Economics* 39, 321–351.

McCracken, Michael W., 2007, Asymptotics for out of sample tests of granger causality, *Journal of Econometrics* 140, 719–752.

Menkhoff, Lukas, and Maik Schmeling, 2010, Whose trades convey information? evidence from a cross-section of traders, *Journal of Financial Markets* 13, 101–128.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Odean, Terrance, 1998, Are investors reluctant to realize their losses?, *The Journal of Finance* 53, 1775–1798.

Payne, Richard, 2003, Informed trade in spot foreign exchange markets: an empirical investigation, *Journal of International Economics* 61, 307–329.

Rime, Dagfinn, Lucio Sarno, and Elvira Sojli, 2010, Exchange rate forecasting, order flow and macroeconomic information, *Journal of International Economics* 80, 72–88.

Roll, Richard, 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *The Journal of Finance* 39, 1127–1139.

Chapter 4

Duration Dependence, Behavioral Restrictions, and the Market Timing Ability of Commodity Trading Advisors[1]

International Review of Finance
In Press

4.1 Introduction

In general, the value potentially added through active management can stem from one or two sources. First, there is the traditional security selection, i.e. the ability to add value by selecting securities that subsequently outperform. Second, managers could also add value by successfully anticipating market trends and reacting to these trends by entering or exiting the market accordingly. This is referred to as market timing ability and has received considerable attention over the last two decades.

[1]This chapter is based on joint work with Michael Frömmel (Ghent University) and Alexander Mende (RPM Risk & Portfolio Management AB).

However, empirical evidence on whether managers do in fact add value through one or both approaches is mixed. One of the first prominent studies on mutual fund performance is Sharpe (1966). He finds no evidence of excess performance for funds compared to the DJIA over the period 1954-1965. This result is confirmed by Jensen (1968), who shows that the average 'alpha' of mutual funds in his dataset is not significantly different from zero. Subsequent evidence is more mixed, but seems to gravitate to the null hypothesis of no significant outperformance by mutual funds. While 'alpha' captures security selection, other studies focus on fund managers' market timing ability, i.e. the ability to adjust ones market exposure in anticipation of future (stock) market movements. The majority of these studies finds no (or sometimes even negative) market timing ability for mutual funds (see e.g. Admati et al., 1986; Becker et al., 1999; Ferson and Schadt, 1996; Henriksson and Merton, 1981; Jensen, 1972; Lehmann and Modest, 1987; Merton, 1981, Kao et al., 1998).

As such, the consensus for mutual funds seems to emerge that mutual fund managers, on average, add little value for investors. To some extent, fees charged by these funds seem to explain most of the lack of performance: Many studies, most recently Fama and French (2010), find that funds' gross returns outperform the market, while the net-of-fee returns do not. This suggests that fund managers are capturing the outperformance through fees.

Evidence for market timing among hedge funds is also mixed, although more recent work indicates some market timing skill for these managers. Whereas Fung et al. (2002) do not find evidence for market timing ability among hedge funds Chen et al. (2010) study a sample of self-declared market timing hedge funds and find evidence of market timing ability. Chen (2007), who examines the timing ability of hedge funds with regard to their focus markets, also finds evidence that a number of categories of hedge funds (CTAs and Global Macro) can successfully time certain asset markets. Finally, Kazemi and Li's (2009) findings suggest that CTAs generate their returns mostly from successful market timing.

However, whereas early studies use monthly returns to test for timing ability, more recent studies such as Bollen and Busse (2001) and Jiang et al. (2007) who use daily data come to more encouraging conclusions about managers' market timing abilities. These results provide evidence that confirm the findings by Goetzmann et al. (2000) that the use of daily data appears to increase the power of the market timing models to detect market timing ability. Chance and Hemler (2001) analyze daily explicit recommendations by market participants and also find evidence of market timing ability. Results in both papers further suggest that, when monthly data is used, the evidence of positive market timing ability disappears.

One major drawback in applying existing market timing models to monthly data is that the researcher implicitly assumes that the trading frequency is also monthly. Goetzmann et al. (2000) are the first to point out this behavioral restriction. The authors propose an adjustment that assumes daily timing but that does not require collecting daily returns. Nevertheless, they note that applying market timing models directly to daily data is preferable. However, the application to daily data creates a potential conflict: Standard tests for market timing (Treynor and Mazuy, 1966; Henriksson and Merton, 1981) use the market's excess return as benchmark for market timing. While this might be a reasonable assumption at lower frequencies, for daily observations it is probably inconsistent with managers' actual timing practices. Both a lack of predictability in daily returns and high transaction costs make such an approach improbable for most funds. Instead, portfolio managers rather think in trends (Menkhoff, 2010). We therefore relax this somewhat restrictive behavioral assumption that is implicit in the application of market timing models on daily data. Instead we use ex-post classified trends as benchmarks.[2] Assuming trend following behavior is particularly justified for CTAs (Fung and Hsieh, 2001). CTAs manage client assets and take long or short positions in highly liquid equity, fixed-income, foreign exchange, metals, and commodity futures markets. Thus, CTAs follow directional strategies and are often described as trend following. Because of CTAs' similarities to hedge funds, they are usually considered a hedge fund category.

Our contribution to the existing literature is twofold. First, we adapt the original Henriksson and Merton market timing model in a way that makes it more realistic and avoids imposing a particular timing frequency. In particular, we replace the 'periodic' timing decision based on monthly or daily excess returns with a definition of timing that depends on (cumulative) past price changes. Obviously, our adjustment also constitutes a re-specification of the market timing definition. Chen and Knez (1996) note, that any performance evaluation is generally arbitrary, a notion that is strongly related to benchmark selection. This also applies to the choice of the benchmark for market timing tests. Our definition of market tim-

[2]The fact that we use an ex-post trend decomposition model does not cause methodological problems, since we do not model managers' decision process. Insofar we are in line with standard market timing models which also rely on ex-post realized market returns. Furthermore, and again in analogy with standard market timing tests, it does not matter whether the detected trends are deterministic or stochastic.

ing differs from the existing excess return-based definition and will lead to different conclusions as to whether CTAs have timing ability. Our alternative definition strongly follows a strand of literature that focuses on formalizing 'bull' and 'bear' states in financial markets using peaks and troughs (see Lunde and Timmermann, 2004; Harding and Pagan, 2002; Pagan and Sossounov, 2003). If successful market timing means successfully timing bull and bear market states, using such a definition provides a natural and meaningful extension of existing market timing models.

Second, we extend the literature on market timing abilities of CTAs using a proprietary dataset of realized audited daily returns of CTAs between 1994 and 2012. Since we use realized instead of reported returns our dataset does not suffer from survivorship bias, backfill bias, or selection bias. Such biases can be meaningful. For example, Bhardwaj et al. (2014) report that the combined backfill and survivorship bias in public hedge fund databases sum up to approximately 7.8% annualized. Furthermore, since the returns we employ are not manipulated, they cannot be a smoothed version of the true realized returns. Spurious serial correlation that results from such smoothing can yield misleading performance statistics (see Getmansky et al., 2004; Agarwal et al., 2011). As the dataset covers the period 1994-2012, it includes the recent financial crisis as well.

The paper proceeds as follows. Section 4.2 describes the methodology including the benchmark model by Henriksson and Merton (1981) and our adaption of the model. Section 4.3 presents the dataset. Section 4.4 discusses our empirical results and conducts a number of robustness checks. Finally, in Section 4.5, we summarize and conclude.

4.2 Methodology

Starting point is the model proposed by Henriksson and Merton (1981) (henceforth HM model). This model assumes that the fund manager allocates capital between a risk-free asset and equities based on a forecast of the market excess return in the next period. To test a manager's market timing ability, the model tests whether the fund's market beta is higher during up-markets than down-markets. To apply the model to data on hedge funds active in multiple markets, we need to extend the approach to a multifactor version of the HM model (see Aragon, 2005; Chen (2007)):

$$r_{p,t} = \alpha + \sum_{m=1}^{M} \beta_m r_{m,t} + \sum_{m=1}^{M} \gamma_m D_{m,t} \cdot r_{m,t} + \lambda_p + \mu_{p,t} \tag{4.1}$$

where λ_p is the time-invariant firm effect of fund p,$r_{p,t}$ is the excess return of fund p at time t,$r_{m,t}$ is the excess return in market m, and $\mu_{p,t}$ is the error term. In the original HM model $D_{m,t}$ is an indicator variable that takes the value 1 if $r_{m,t} > 0$ and zero otherwise. The coefficient γ_m measures the difference in betas in down- vs. up-markets. γ_m will be significantly positive for a manager who successfully times market m. The HM model does not allow the manager to vary her exposure in any but the most restrictive way. In particular, depending on her forecast, the manager chooses two levels of β. While this assumption can be considered restrictive or inappropriate in the case of mutual funds, the model adequately describes the trading strategy of certain types of hedge funds and CTAs in particular. CTAs either buy or sell futures contracts in a particular market, which is arguably the type of systematic risk variation assumed under the HM model.

Previous research on the timing ability of hedge funds relied on constructing equal-weighted portfolios (see Chen, 2007; Kazemi and Li, 2009) to test for market timing ability among hedge funds. However, since we have a panel of daily CTA observations, we have considerably more degrees of freedom than previous work which commonly employed monthly data. Therefore, a panel approach is more appropriate as it allows more accurate inference of the model parameters.

We estimate the model using fixed effects for each fund. This estimation approach allows us to account for managers' fixed effects that are unrelated to market timing ability. For example, some funds in the sample could be persistently more profitable for reasons that we do not observe. At the same time, we also cluster the standard errors by manager because, although the fixed effect dummies handle the fund effects, the dummies will not handle some other relevant forms of correlated errors (Thompson, 2011).

In addition to manager fixed effects, time fixed effects might also be present. Given that the managers are actively trading the same futures markets, it is unlikely that the observations on the different managers within every time period are not correlated. As such, the dataset can be expected to contain time effects beyond those we are interested in. Moreover, these time effects are probably not fixed. We can imagine that some CTAs perform better than others, depending on the particular market environment. However, while our panel is extremely unbalanced, any bias present in the standard errors due to time effects is likely to disappear since we have a lot of observations along this dimension. This also explains why we cluster on the less numerous (i.e. manager) dimension.

The construction of the dummy $D_{m,t}$ is a key component of the HM model. The HM model, however, imposes a timing frequency that matches the return frequency used to estimate the model. As we have already discussed in the introduction, performance evaluation is generally arbitrary (Chen and Knez, 1996). This observation also applies to the benchmark of what constitutes proper market timing. Consequently, the alternative definition of market timing we put forth below differs from the above excess return-based definition and might therefore lead to different conclusions as to whether CTAs have timing ability under either definition. Our definition borrows extensively from recent literature that focuses on formalizing bull and bear market states in financial markets using peaks and troughs (see Lunde and Timmermann, 2004; Harding and Pagan, 2002; Pagan and Sossounov, 2003). If by successful market timing investors mean successfully timing bull and bear states, then our definition provides a natural extension of existing market timing models. In addition, such a definition is in line with the observation that market professionals think in terms of trends, rather than in terms of excess returns (Menkhoff, 2010).

Therefore, a dummy variable based on a trend identification scheme seems to be a reasonable alternative to assuming that funds in general, and CTAs in particular, make predictions only about the next period's excess return. This might be especially relevant when evaluating funds' performance over very short time horizons. However, an application of existing market timing models on daily data implies exactly that. Temporary drops or increases in asset prices over several days can be expected to be short-lived and might only induce partial adjustments or no adjustment at all. This is especially the case if we consider transaction costs, which can make daily adjustments based on daily forecasts of excess returns costly.

We identify trends in asset markets by drawing on the academic literature that proposes methods to determine bull and bear states in stock markets. This literature offers both parametric and nonparametric approaches.[3] We rely on a threshold filter recently suggested by Lunde

[3]The most popular parametric approach imposes a Markov-switching model (Hamilton, 1990) that allows for two regimes, booms and busts. Examples of applications of

and Timmermann (2004), which is described in the Appendix. This filter has the advantage that it allows for duration dependence and does not impose a phase length.[4] The threshold filter proposed by Lunde and Timmermann identifies bull and bear markets based on a minimum price change ('threshold') since the last peak or trough. Whereas an excess return-based measure will classify a given period of negative price movements as a bear market, the Lunde and Timmermann filter will not as long as the drop does not exceed a certain threshold.

The drawback of this rules-based method is that we need to specify the thresholds that define bull and bear markets.[5] Lunde and Timmermann (2004) suggest such thresholds only for equity markets, based on figures for bull and bear markets commonly reported in the financial press. However, since we also want to explore CTAs' market timing ability in other asset markets, we first have to derive additional thresholds. Since previous literature has not yet proposed a method to come up with such thresholds, we employ an approach inspired by the work of Wegscheider (1994). This method aims to identify trends, store their magnitude, and subsequently remove them in an iterative way until all trends are identified. The advantage of this algorithm is that, rather than imposing some arbitrary structure on the data, it focuses on the specific features of the original data series to come up with thresholds. What we obtain is a set of trends, starting from very small trends that last just one day to trends that last several months. This makes it an ideal tool to derive appropriate thresholds for the Lunde and Timmermann filter. We describe the algorithm in detail in the Appendix.

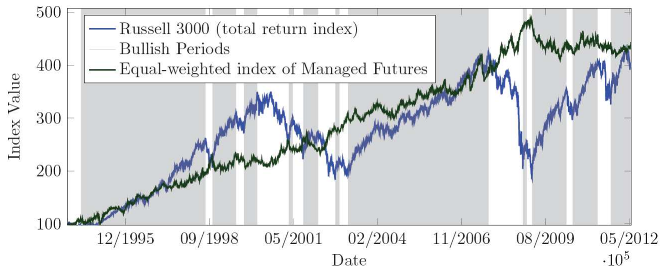
---

this approach in the context of stock markets are Maheu and McCurdy (2000) and Chen (2009). Nonparametric approaches rely on filters or dating algorithms that locate turning points (peaks and troughs) corresponding to local maxima and minima of the financial series. Pagan and Sossounov (2003) modify the algorithm developed by Bry and Boschan (1971) using definitions on the duration of bull and bear markets found in financial press. Lunde and Timmermann (2004) construct a filter that identifies bull and bear markets based on a minimum price change since the most recent peak or trough.

[4]Duration dependence means that "bull and bear hazard rates – that is, the probability that a bull or bear market terminates in the next period – depend on the age of the market" (Lunde and Timmermann, 2004, p253).

[5]We want to avoid misclassification through imposing restrictions on the timing frequency. Therefore, we cannot make use of the algorithm of Bry and Boschan (1971), since this approach requires choosing the phase length.

Figure 4.1: Evolution CTA Index



4.3 Data

We carry out the empirical analysis using a proprietary dataset of daily returns on 33 CTAs for the period January 1994 to May 2012. The data is provided by a Swedish CTA specialist and fund of funds manager.[6] We focus on CTAs because CTAs can be considered a hedge fund category that actively attempts to perform market timing. Returns are raw returns in that they exclude manager fees and trading commissions and, thus, provide an unbiased account of realized returns. The dataset does not suffer from most of conventional biases found in public data bases due to voluntarily reporting by funds. In particular, the data base does not suffer from survivorship bias, backfill bias, or selection bias. Furthermore, since the returns are not reported returns, they cannot be a smoothed version of the true realized returns. This is important, since the spurious serial correlation resulting from such smoothing yields misleading performance statistics (see Getmansky et al., 2004).

In the sample of CTA funds, 26 are active across different asset markets ('diversified'), four funds trade exclusively in financials, and three funds invest only in commodity futures. The time frame covers a variety of market conditions including several financial crises. During 1994-2012, markets have experienced pronounced directional moves. This makes the sample period ideal to test for market timing ability. In Figure 4.1, we plot the performance of an equally-weighted index of the CTAs' returns and compare it to the Russell 3000 Total Return Index. Shaded areas correspond to bull market phases (as defined below).

To test for market timing ability for the main asset classes CTAs invest

[6]We do not identify the names of the CTAs in the dataset.

in, we use daily observations for the following market indices: the Russell 3000 for equities, Barclays US Aggregate Bond Index, the S&P GSCI Agricultural Commodities Spot Index, the S&P GSCI Energy Spot Index, the S&P GSCI All Metals Spot Index and the Fed's trade-weighted US Dollar Index.[7] These market indices encompass the different asset classes managed futures managers are active in.

In particular, we follow Fung and Hsieh (1997) and Agarwal and Naik (2000; 2004). With some variation regarding particular indices used, these authors consider a broad US stock market index, a US bond index, the Fed's trade-weighted US Dollar index, and the Goldman Sachs Commodity Index (GSCI) as proxy for markets that hedge funds have exposure to. We deviate from the above studies in that we break down the Goldman Sachs Commodity Index in its various components. We do this because CTAs have historically been active mostly in commodity markets for which futures were first available. They might therefore have skills particularly in these markets. The pairwise correlations between the indices, reported in Panel C of Table 4.1, are relatively low. This indicates that the co-movement on a daily basis between the different markets is generally limited. The pairwise correlation is highest among commodity indices but it is still sufficiently low to justify a separate treatment.

4.4 Results

We start by applying the algorithm proposed by Wegscheider (1994) to the various markets. Once we have identified the trends in different markets, we select the 99 percentile of trends found. In Table 4.2 we report the results of the approach. Following Lunde and Timmermann (2004) we allow for different cut-off values in the case of upward and downward trends. This allows us to account for a positive drift in certain asset classes and potential asymmetries in up and down trends.

For the equity market index, our results indicate that the top 1 percentile of upward trends exceeds 19.04% while the corresponding value for downward trends is only -10.22%. These values are close to the ones reported in the financial press and the ones Lunde and Timmermann (2004) use (20% and 10% for bull and bear markets, respectively). Cut-off values for the other asset classes differ considerably from the values for stock markets. For example, large trends in the bond market that are similar in

[7] In line with Lunde and Timmermann we use daily *price* indices to identify trends in the different markets.

Table 4.1: Descriptive Statistics of the Dataset

Panel A: Summary Statistics CTAs

	Mean	Min	P25	P50	P75	Max
Return	0.0150%	-0.0743%	0.0001%	0.0161%	0.0360%	0.0703%
Standard deviation	0.72%	0.21%	0.54%	0.79%	0.89%	1.19%
Age	3.9	0.3	1.5	2.7	4.9	13.9
Skewness	-0.152	-2.006	-0.468	-0.221	0.154	3.523
Kurtosis	9.058	3.258	5.582	6.798	9.259	66.580

Panel B: Summary Statistics Factors

Market	Index	Mean return	Standard deviation	Min	Max
EQUIT	Russell 3000 TR	0.01%	0.54%	-4.23%	4.72%
BOND	Barclays US Aggr. Bond	0.00%	0.11%	-0.77%	0.59%
AGRI	S&P GSCI Agri. Commodity	0.00%	0.52%	-3.32%	3.11%
ENER	S&P GSCI Energy Spot	0.02%	0.84%	-6.25%	4.26%
METAL	S&P GSCI All Metals Spot	0.01%	0.50%	-3.11%	2.90%
CUR	Fed's Trade-Weighted USD	0.00%	0.14%	-1.25%	1.24%

Panel C: Correlation Market Indices

Market	EQUIT	BOND	AGRI	ENERGY	METAL	CUR
EQUIT	1.00					
BOND	-0.12	1.00				
AGRI	0.15	-0.09	1.00			
ENERGY	0.15	-0.07	0.28	1.00		
METAL	0.21	-0.09	0.31	0.30	1.00	
CUR	-0.13	-0.02	-0.23	-0.21	-0.39	1.00

This table reports summary statistics for the set of CTAs and the factors used in the multifactor approach.

Table 4.2: Results Identification Bull and Bear Markets

	Cut-off values			
	Upward trends	Standard deviation	Downward trends	Standard deviation
Russell 3000 TR Index	19.04%	6.08%	-10.22%	2.81%
Barclays US Aggregate Bond Index	2.56%	0.58%	-4.63%	0.99%
GSCI Agricultural Commodities Index	13.85%	2.68%	-19.58%	4.23%
GSCI Energy Spot Index	23.90%	9.18%	-19.71%	4.62%
GSCI All Metals Spot Index	10.61%	2.36%	-16.01%	3.60%
Fed's Trade-weighted USD Index	4.07%	1.23%	-4.59%	4.44%

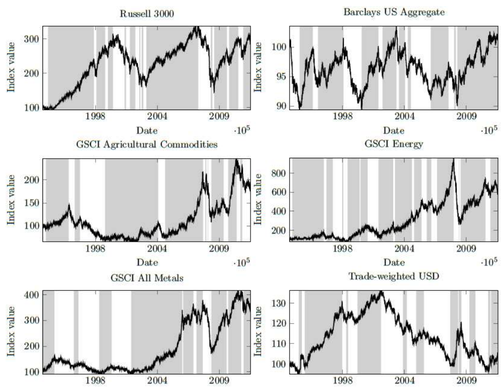
	Panel B: Concordance Index			
Market	EQUIT	BOND	AGRI	ENER
EQUIT	1			
BOND	0.576	1.000		
AGRI	0.644	0.649	1.000	
ENER	0.676	0.640	0.663	1.000
METAL	0.560	0.570	0.468	0.555
CUR	0.474	0.386	0.242	0.421

The concordance index measures the fraction of the time the cycles are in the same state. If the index is unity, trends in both markets are exactly pro-cyclical, while a value of zero indicates that they are perfectly countercyclical. For two series y_t and x_t and a sample size of T, the index can be calculated as:

$$\hat{I} = \frac{1}{T}\left[\sum_{t=1}^{T} S_{x,t}S_{y,t} + \sum_{t=1}^{T}(1 - S_{x,t}) \cdot (1 - S_{y,t})\right]$$

where $S_{x,t}$ and $S_{y,t}$ are dummies that equal 1 in the case of an upward trends and zero otherwise.

Figure 4.2: Bull and Bear Markets Identified



frequency only exceed 2.56% for up markets and -4.63% for down markets. The largest trends are reported for the S&P GSCI energy market, with upward trends of over 23.90% and downward trends exceeding -19.71%.[8] The results support our view that a separate trend classification for every asset class is necessary. It would prove unrealistic to generalize the equity-based thresholds from the financial press to other asset classes.

Based on the thresholds derived above we can employ the filter suggested by Lunde and Timmermann to obtain a classification of the markets into bull and bear market periods. The results are reported in Figure 4.2 with bull markets periods shaded grey. Obviously, the filter identifies major market events such as the dotcom bubble, the bull market between 2003 and 2006 for stocks. It also captures major surges in agricultural commodity, energy, and metal prices. To measure the degree of co-movement between the trends, we employ the concordance index, proposed by Harding and Pagan (2002). The results, reported in Table 4.2, show that markets are in the same market state about half to two-thirds of the time, depending on the markets under consideration. Of course, this does not necessarily mean that they start and end at the same time. Two markets might be trending upwards two-thirds of the time, but both market might nevertheless experience bear markets at different points in time.

[8] A similar analysis was performed using the S&P 500 as the equity index, yielding 19.00% and -10.80%, respectively.

In unreported tests, we test for the presence of duration dependence given our classification. In particular, we apply the tests by Shapiro and Wilk (1972), Brain and Shapiro (1983), and Ohn et al. (2004) for duration dependence. All tests indicate statistically significant duration dependence in both the equity and currency market. For agricultural commodities, only the result from the Shapiro and Wilk (1972) test is significant at the 10% level. These results confirm our view that duration dependence plays a role in an number of markets under consideration and that the threshold filter of Lunde and Timmermann should be preferred.

4.4.1 Market Timing Ability

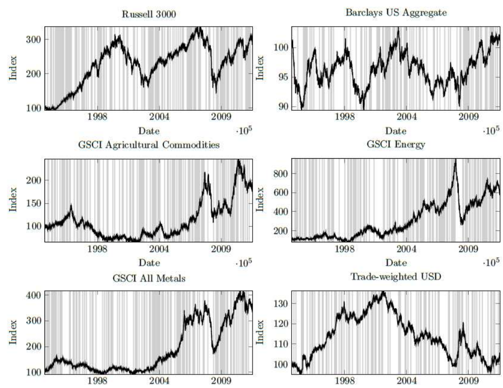
We now turn to our main analysis, testing whether CTAs are able to successfully time the bull and bear markets identified above. We report the results for the main regressions in Table 4.3.

The first set of regression results, corresponding to our baseline model outlined in Section 4.2, suggests that CTAs exhibit market timing ability in all of the markets considered. All the interaction terms measuring timing ability are highly significant and show the expected sign.[9] The intercept, which is the average value of the manager fixed effects, is significantly negative. Although returns of the funds are before fees and transaction costs, nevertheless, they reflect implicit transaction costs. In particular, the negative coefficient on the intercept likely reflects bid-ask spreads.

Turning to the economic significance of the timing coefficients reported in Table 4.3, we see that the magnitude of the observed market timing is meaningful. For example, a 1% increase in bond markets when bond markets exhibit a positive trend is associated, on average, with a 1.28% (0.313% + 0.967%) return to the fund. When bond markets are declining, however, the funds' returns are only associated with a decrease by 0.313% on average for every 1% decrease in bonds. In other words, funds tend to exhibit a significantly positive beta to bond markets during up-markets, but an insignificant beta during down-markets. Similarly, all else equal, a 1% increase in the trade-weighted US dollar index during up-trends is

[9] We note that these results do not allow us to infer the extent to which a manager anticipates trends in a particular asset class on a stand-alone basis. In particular, managers' timing ability in one market can be the result of successfully anticipating the trends in other markets. The high degree of overlapping in market states, as evidenced by the concordance index calculated in Table 4.2, makes this a likely possibility. For example, we can imagine that if a manager expects a strong reversal in the stock market, she will use that information to adjust her exposure to, say, energy markets.

Figure 4.3: (Monthly) Excess Return-based Classification



associated to a 0.389% (0.938%-0.549%) increase in the funds' returns, whereas the funds seem to gain 0.549% for every 1% decrease in the index during down-markets.

Next, we contrast these findings with the results obtained for two existing models. First, we apply the HM model to daily data, where the dummy variable is one when the excess return for the *month* is positive and zero otherwise. A visual illustration of the classification that results from the HM model is shown in Figure 4.3.Clearly, this classification leads to a more dispersed set of up- and down market periods.

Column (2) of Table 4.3 reports the results when we employ this definition of bull and bear markets. The coefficients of the timing variables suggest that in this case, too, i.e. CTAs exhibit timing ability in four out of six markets considered. This result reveals that also under the traditional definition of market timing ability, CTAs show clear evidence of market timing skill.

Finally, we also consider the daily version of the HM model suggested by Bollen and Busse (2001), where instead of using monthly excess returns, we look at dailies. In days where the excess return is positive, the dummy is one, while it is zero otherwise. This approach is commonly followed when researchers have access to daily data. The results, reported in column (3) of Table 4.3, are striking. The estimates suggest that when using this definition of market timing, CTAs do not exhibit any timing skill. On the contrary, we find evidence of significantly negative timing

Table 4.3: Market Timing Ability of CTAs

	(1)	(2)	(3)
Equities	-0.161***	-0.121***	-0.145***
	(0.0388)	(0.0316)	(0.0227)
Equities $\cdot D_{1,t}$	0.158***	0.0454	0.0649
	(0.0411)	(0.0311)	(0.0522)
Bonds	0.313	0.700***	1.188***
	(0.190)	(0.167)	(0.183)
Bonds $\cdot D_{2,t}$	0.967***	0.563***	-0.527***
	(0.245)	(0.164)	(0.127)
Agri. Commodities	-0.0237	-0.0480**	-0.00113
	(0.0223)	(0.0181)	(0.0223)
Agri. Commodities $\cdot D_{3,t}$	0.0714**	0.114***	0.0277
	(0.0278)	(0.0232)	(0.0266)
Energy	-0.0284	0.0304*	0.0635***
	(0.0194)	(0.0178)	(0.0188)
Energy $\cdot D_{4,t}$	0.157***	0.0682***	-0.0134
	(0.0276)	(0.0220)	(0.0169)
Metals	0.0349	0.156***	0.214***
	(0.0392)	(0.0279)	(0.0358)
Metals $\cdot D_{5,t}$	0.146***	-0.0115	-0.144***
	(0.0415)	(0.0234)	(0.0320)
Currencies	-0.549***	-0.204*	-0.119
	(0.104)	(0.120)	(0.120)
Currencies $\cdot D_{6,t}$	0.938***	0.209*	-0.00406
	(0.132)	(0.118)	(0.108)
Constant	-0.000169***	3.36e-05	0.000578***
	(6.04e-05)	(4.27e-05)	(0.000146)
Observations	32,450	32,450	32,450
Adj. R-squared	0.070	0.044	0.040
Number of funds	33	33	33

This Table reports the results for Eq (1), using different definitions for the market timing dummies. Column (1) reports the results for the specification that employs a bull- and bear market definition using the approach of Lunde and Timmermann (2004). Column (2) reports the results using the definition proposed by Henriksson and Merton (1981). Finally, column (3) reports the results using the specification of Bollen and Busse (2001).

Cluster-robust standard errors in parentheses. Significance at 1%, 5%, and 10% level indicated by ***, **, and *, respectively.

skill. The reason for this result might relate to the behavioral restriction that is implicit in a direct application of the HM model to daily data. Such an application of the model implicitly assumes that market timing is executed on a daily basis, but as mentioned above this restriction seems too binding for the funds under consideration.

4.4.2 Robustness Checks

To verify whether our proposed approach, i.e. analyzing funds' market timing ability in terms of trends rather than in terms of excess returns, indeed adds value, we perform a number of robustness checks.

4.4.2.1 Correlation across Time

To test the significance of the results, we have ignored the potential impact of correlation across time. We cluster on the less numerous (i.e. by firm) dimension following the suggestions of Petersen (2008) and Thompson (2011). In particular, if the time dimension is considerably larger than the firm dimension, the bias due to correlation can be expected to disappear as long as one (single)-clusters on the less numerous dimension. It may nevertheless be instructive to cluster by time as well, since the regressors vary by time but not by firm.

To this end, we perform a number of robustness checks to test whether our results are robust to correlation across time. First, we include the regression results where we include time fixed effects. At the same time, we still cluster the standard errors by fund. This is one way of simultaneously handling firm and time fixed effects, although there are also limitations to such an approach (see Thompson, 2011). The first column of Table 4.4 reports the results, where we omit the dummy for 1994 to serve as reference category. We find that our results are robust to time fixed effects.

Next, we also report the results where standard errors are clustered by time and clustered both by time and by firm (two-way clustering). Clustering simultaneously by time and firm follows the work of Thompson (2011) and Petersen (2009). Column (2) and (3) of Table 4.4 report the results for clustering by time and two-way clustering, respectively. We find that our results are robust to clustering along both dimensions.

Table 4.4: Robustness to Correlation across Time

	(1)	(2)	(3)
Equities	-0.1603***	-0.1589***	-0.1589***
	(0.041)	(0.034)	(0.047)
Equities $\cdot D_{1,t}$	0.1669***	0.1630***	0.1630***
	(0.045)	(0.046)	(0.057)
Bonds	0.3102	0.3041*	0.3041
	(0.190)	(0.171)	(0.234)
Bonds $\cdot D_{2,t}$	0.9702***	0.9789***	0.9789***
	(0.248)	(0.222)	(0.307)
Agri. Commodities	-0.0276	-0.0256	-0.0256
	(0.024)	(0.038)	(0.040)
Agri. Commodities $\cdot D_{3,t}$	0.0777**	0.0760*	0.0760
	(0.029)	(0.045)	(0.048)
Energy	-0.0265	-0.0257	-0.0257
	(0.020)	(0.025)	(0.030)
Energy $\cdot D_{4,t}$	0.1565***	0.1538***	0.1538***
	(0.029)	(0.030)	(0.039)
Metals	0.0451	0.0438	0.0438
	(0.042)	(0.044)	(0.055)
Metals $\cdot D_{5,t}$	0.1553***	0.1551***	0.1551***
	(0.044)	(0.049)	(0.060)
Currencies	-0.5649***	-0.5456***	-0.5456***
	(0.111)	(0.119)	(0.153)
Currencies $\cdot D_{6,t}$	1.0043***	0.9638***	0.9638***
	(0.145)	(0.166)	(0.200)
Constant	-0.000295	-0.000275***	-0.000275***
	(0.001)	(0.000)	(0.000)
Time Fixed Effects	Yes		
Observations	32,450	32,450	32,450
Adj. R-squared	0.069	0.067	0.067

This Table reports the results for a robustness checks where we test the robustness of the specification in the first column of Table 4.3 for correlation across time. Column (1) reports the results for a specification where we include time fixed effects. Column (2) reports the results when we cluster by time. In column (3), we report the results from clustering both by time and by firm (two-way clustering). Standard errors in parentheses. Significance at 1%, 5%, and 10% level indicated by ***, **, and *, respectively.

4.4.2.2 Thresholds

We also test whether the baseline result in Section 4.4 is robust to the thresholds used. While the thresholds for the stock market are close to the ones proposed by Lunde and Timmermann (2004), the other thresholds are not well-established yet. Therefore, we redo the analysis with bigger (smaller) thresholds. In particular, we increase (decrease) the absolute value of the thresholds by one standard deviation to look at whether our results still hold for somewhat higher (smaller) trends. The results, reported in column (1) and (2) Table 4.5, suggest that our baseline results are only to a minor extent driven by the particular set of thresholds obtained in Section 4.4. Especially, CTAs seem to be successful at timing the larger trends in currencies, since for the smaller trends the managers show negative timing ability. Similarly, the funds do not show timing ability for the very large trends in agricultural commodities. Nevertheless, the explanatory power of our model seems to be increasing with the size of the trend. This suggests that CTAs' market timing ability takes the form of successfully timing the larger trends in the different markets.

4.4.2.3 Microstructure issues

The use of daily fund data might lead to microstructure related issues such as possible thin or nonsynchronous trading and stale pricing (Scholes and Williams, 1977). It is unlikely that our results are driven by such issues, given the nature of the futures markets CTAs trade in. Nevertheless, we re-estimate our baseline model but include lagged values for the market factors (Dimson, 1979). In that case, the model changes to:

$$r_{p,t} = \alpha + \sum_{m=1}^{M} \beta_{1,m} r_{m,t} + \sum_{m=1}^{M} \beta_{2,m} r_{m,t-1} + \sum_{m=1}^{M} \gamma_m D_{m,t} \cdot r_{m,t} + \lambda_p + \mu_{p,t} \tag{4.2}$$

The results, reported in column (3) of Table 4.5, show that these concerns are unwarranted. Including lagged market factors does not materially impact results for the variables of interest.

4.4.2.4 Conditional Performance

To ensure that funds indeed add value in successfully timing markets, we also investigate the performance *conditional* on public information. This approach, suggested by Ferson and Schadt (1996), is motivated from the

Table 4.5: Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
Equities	-0.162***	-0.173***	-0.173***	-0.120***	-0.230***	-0.199***
	(0.0356)	(0.0397)	(0.0402)	(0.0393)	(0.0363)	(0.0456)
Equities $\cdot D_{1,t}$	0.187***	0.196***	0.165***	0.191***	0.206***	0.135**
	(0.0421)	(0.0434)	(0.0413)	(0.0375)	(0.0685)	(0.0643)
Bonds	0.109	0.269	0.317	0.0506	0.110	0.260
	(0.220)	(0.175)	(0.192)	(0.219)	(0.184)	(0.216)
Bonds $\cdot D_{2,t}$	1.204***	1.102***	0.984***	0.747***	0.241	0.0474
	(0.289)	(0.186)	(0.251)	(0.262)	(0.228)	(0.302)
Agri. Com	0.0237	-0.0975***	-0.0181	-0.00107	0.0692	0.00185
	(0.0236)	(0.0212)	(0.0219)	(0.0272)	(0.0438)	(0.0296)
Agri. Com $\cdot D_{3,t}$	-0.00011	0.200***	0.0640**	0.0769**	-0.0372	0.0673
	(0.0267)	(0.0263)	(0.0273)	(0.0302)	(0.0458)	(0.0442)
Energy	-0.0166	0.00772	-0.0239	0.0332**	0.0348	0.0215
	(0.0184)	(0.0150)	(0.0185)	(0.0147)	(0.0317)	(0.0234)
Energy $\cdot D_{4,t}$	0.151***	0.105***	0.153***	0.109***	0.0186	0.0356
	(0.0280)	(0.0205)	(0.0267)	(0.0208)	(0.0377)	(0.0342)
Metals	0.0113	0.0454	0.0472	0.0942**	-0.135	-0.0706
	(0.0459)	(0.0390)	(0.0391)	(0.0436)	(0.0813)	(0.0445)
Metals $\cdot D_{5,t}$	0.175***	0.148***	0.137***	0.0921**	0.164**	0.126*
	(0.0510)	(0.0433)	(0.0412)	(0.0407)	(0.0800)	(0.0706)
Currencies	-0.809***	0.00974	-0.580***	-0.451***	-0.411***	-0.395**
	(0.118)	(0.115)	(0.110)	(0.134)	(0.134)	(0.148)
Currencies $\cdot D_{6,t}$	1.277***	-0.334**	0.912***	0.860***	0.472***	0.427**
	(0.144)	(0.123)	(0.130)	(0.138)	(0.134)	(0.166)
Constant	-0.0001**	-0.0001**	-0.0002**	-0.00014**	-0.0014	-0.0060*
	(5.57e-05)	(5.91e-05)	(6.13e-05)	(5.33e-05)	(0.0019)	(0.0034)
Time Fixed Effects			Yes			
Controls for macro-economic information				Yes		
	32,450	32,450	32,449	32,449	1,486	1,486
Observations	0.079	0.062	0.073	0.106	0.088	0.064
Adj. R-squared	33	33	33	33	33	33

This Table reports the results for a number of robustness checks. In column (1) and (2) we test the robustness of the results to higher and lower thresholds, respectively. In column (3) we include lagged market factors to account for potential microstructure issues. In column (4) we control for publicly available information, following Ferson and Schadt (1996). Column (5) and (6) report the results from estimating the bull and bear market and the Henriksson-Merton specification the using monthly data.
Cluster-robust standard errors in parentheses. Significance at 1%, 5%, and 10% level indicated by ***, **, and *, respectively.

idea that profitable trading strategies relying on public information should not yield superior performance. To estimate this model, we make use of four different macroeconomic variables to control for publicly available information: a dividend yield, a liquidity premium, a default risk premium, and the risk-free rate.[10] All four variables are constructed using daily data. Following Ferson and Schadt (1996), the variables are demeaned and their lagged values are interacted with the market factors.

The model takes the following form:

$$\begin{aligned} r_{p,t} \quad = \quad & \alpha + \textstyle\sum_{m=1}^{M} \beta_{1,m} r_{m,t} + \sum_{m=1}^{M} \beta_{2,m} r_{m,t-1} + \sum_{m=1}^{M} \gamma_m D_{m,t} \cdot r_{m,t} \\ & + \textstyle\sum_{m=1}^{M} \sum_{n=1}^{4} \emptyset_{m,n} r_{m,t} \cdot c_{n,t-1} + \lambda_p + \mu_{p,t} \end{aligned} \tag{4.3}$$

where $c_{n,t-1}$ represent the lagged and demeaned macroeconomic variables. These interaction terms pick up the movements through time of the conditional betas as they relate to the market indicators. Column (4) of Table 4.5 reports the conditional market timing performance of the CTAs, which suggests that the CTAs' successful time-varying exposure to the different factors cannot be explained by publicly available information. Incidentally, the inclusion of these macro-economic variables also controls for potential common shocks. This specification therefore provides complementary evidence that our results are robust to time fixed effects.

4.4.2.5 Return Frequency

Next, we test the impact of the frequency of the return data on our results. Previous literature commonly relied on monthly data, mainly due to data availability issues. Bollen and Busse (2001) show that evidence of monthly timing ability tends to disappear when daily data is employed. To verify whether our results are also sensitive to the data frequency, we redo the analysis using monthly data. In particular, we redo both the specification bull and bear market specification and the excess return-based specification.

[10]The term spread, which proxies for the liquidity premium, is calculated as the difference between the US Treasury 10 year yield and the (annualised) three-month US T-Bill yield. The latter also serves as the risk-free rate. The quality spread is the difference between the US Corporate Bonds Moody's Seasoned AAA and the US Corporate Bonds Moody's Seasoned BAA rate. The dividend yield is the daily dividend yield of the S&P 500.

Results are shown in columns (5) and (6) of Table 4.5. When employing our baseline model to monthly data, we find the evidence of positive timing ability disappearing for half of the markets under consideration. The results for the original HM model using monthly data yield identical results. The evidence of positive timing ability reported in column (2) of Table 4.3 is no longer present in column (6) of Table 4.5. These results are in line with previous literature and illustrate the importance of using daily data for testing market timing ability.

4.4.2.6 Impact of fees

An analysis of alpha after fees provides another dimension along which we can evaluate the robustness of our results. Such an analysis is relevant since hedge funds' fee structure impacts net-of-fee returns in a non-linear way. This is the case since part of hedge fund managers' compensation is based on performance relative to a high-water mark. To assess the impact of fees, we re-estimate the main specifications in the paper (Table 4.3) using after fee returns. The results are reported in Table 4.6.

We find that our results are robust to the use of net-of-fee returns. The only change that we observe, is a slight drop in the constant. This is consistent with the findings of Kazemi and Li (2009) who note that, since CTAs do not engage in security selections, the slightly negative constants may be the result of fees and transaction costs.

4.4.2.7 Subsample Analysis

Finally, we perform a subsample analysis to investigate how CTAs' market timing ability has evolved over time. We use subsamples defined by events. In particular, we look at the period up to the dotcom crash (1994-1999), the period of the crash and subsequent bull market (2000-2007) and finally the recent financial crisis (2008-2012). We report the results for the three sample periods in Table 4.7.

In general, we find that there has been some time variation in CTAs' timing ability of trends in the different markets under consideration. For the period 1994-1999, CTAs exhibit positive timing ability in markets, although only significantly so in half of the cases. In contrast, while timing ability with regard to equity markets improves considerably during the second sub-period, the results suggest a clear absence of timing ability in agricultural markets. Finally, the period 2008-2012 suggests an overall improvement in the timing ability of CTAs, compared to the previous two

Table 4.6: Market Timing and Net-of-fee Returns

	(1)	(2)	(3)
Equities	-0.1593***	-0.1157***	-0.1388***
	(0.041)	(0.033)	(0.022)
Equities $\cdot D_{1,t}$	0.1637***	0.0418	0.0600
	(0.044)	(0.034)	(0.053)
Bonds	0.3055	0.7086***	1.2023***
	(0.188)	(0.169)	(0.185)
Bonds $\cdot D_{2,t}$	0.9771***	0.5416***	-0.5560***
	(0.244)	(0.174)	(0.129)
Agri. Commodities	-0.0275	-0.0487**	-0.0007
	(0.024)	(0.019)	(0.024)
Agri. Commodities $\cdot D_{3,t}$	0.0795***	0.1173***	0.0291
	(0.029)	(0.025)	(0.028)
Energy	-0.0261	0.0305*	0.0639***
	(0.020)	(0.018)	(0.019)
Energy $\cdot D_{4,t}$	0.1540***	0.0678***	-0.0143
	(0.029)	(0.021)	(0.017)
Metals	0.0420	0.1714***	0.2344***
	(0.042)	(0.030)	(0.041)
Metals $\cdot D_{5,t}$	0.1578***	-0.0117	-0.1545***
	(0.044)	(0.024)	(0.037)
Currencies	-0.5523***	-0.1877	-0.1015
	(0.111)	(0.125)	(0.127)
Currencies $\cdot D_{6,t}$	0.9762***	0.2033	-0.0093
	(0.141)	(0.123)	(0.114)
Constant	-0.000279***	-6.30e-05	0.000524***
	(0.000)	(0.000)	(0.000)
Observations	32,450	32,450	32,450
Adj. R-squared	0.067	0.042	0.038
Number of funds	33	33	33

This Table reports the results for a robustness checks where we re-estimate the specifications in Table 3 using net-of-fee returns, rather than gross returns. Column (1) reports the results based on a bull- and bear markets using the algorithm of Lunde and Timmermann (2004). Column (2) reports the results for the specification that uses the classification of Henriksson and Merton (1981). Finally, column (3) reports the results using the approach of Bollen and Busse (2001).
Standard errors, clustered by fund, in parentheses.
Significance at 1%, 5%, and 10% level indicated by ***, **, and *, respectively

Table 4.7: Subsample Analysis

	1994-1999	2000-2007	2008-2012
Equities	-0.182	-0.286***	-0.0800*
	(0.141)	(0.0418)	(0.0455)
Equities $\cdot D_{1,t}$	0.190	0.334***	0.0670
	(0.143)	(0.0413)	(0.0460)
Bonds	-1.183**	0.198	1.176***
	(0.239)	(0.261)	(0.159)
Bonds $\cdot D_{2,t}$	3.101**	1.004***	-0.0800
	(0.614)	(0.232)	(0.222)
Agri. Commodities	-0.120*	0.116**	0.00374
	(0.0391)	(0.0421)	(0.0217)
Agri. Commodities $\cdot D_{3,t}$	0.412	-0.110**	0.0879***
	(0.227)	(0.0390)	(0.0259)
Energy	-0.0849*	0.0104	-0.0886**
	(0.0295)	(0.0168)	(0.0314)
Energy $\cdot D_{4,t}$	0.316***	0.124***	0.174***
	(0.0339)	(0.0242)	(0.0503)
Metals	-0.206**	0.143**	0.0134
	(0.0388)	(0.0682)	(0.0377)
Metals $\cdot D_{5,t}$	0.331***	0.0459	0.178***
	(0.0504)	(0.0713)	(0.0462)
Currencies	0.0676	-1.004***	-0.213**
	(0.214)	(0.146)	(0.0773)
Currencies $\cdot D_{6,t}$	0.986	1.407***	0.508***
	(0.480)	(0.250)	(0.169)
Constant	-0.000194	-7.47e-05	-0.000337***
	(0.000155)	(4.79e-05)	(0.000112)
Observations	2,724	17,857	11,846
Adj. R-squared	0.119	0.090	0.070
Number of funds	33	33	33

This Table reports the results for a subsample analysis. Cluster-robust standard errors in parentheses. Significance at 1%, 5%, and 10% level indicated by ***, **, and *, respectively.

sub-periods. With the exception of bonds and equities, CTAs seem to have successfully timed the other markets under consideration. This finding is consistent with the clear trends that emerged during the financial crisis. The absence of timing ability in stock and bond markets is consistent with anecdotal evidence that CTAs got whipsawed in these markets following the risk-on/risk-off environment after 2009.

4.5 Concluding Remarks

In this paper we extend the well-established Henriksson-Merton model for market timing by using a less restrictive assumption on managers' objectives. In particular, we assume that the manager attempts to time bull and bear markets, rather than expected excess returns over the next period (i.e. next month or next day). As such, our analysis bridges the literature on bull and bear market identification and tests for market timing ability. Our approach builds on the observation that market professionals think in trends rather than in terms of excess returns.

Since any performance evaluation is generally arbitrary, we test whether market participants succeed in timing the trends we identify using our proposed definition. In particular, we test whether CTAs, a hedge fund category that attempts to profit from trends, are able to successfully time bull and bear periods in the asset classes they are generally active in. Our results suggest that CTAs exhibit market timing ability and are generally able to successfully time trends in financial markets.

Bibliography

Admati, Anat R., Sudipto Bhattacharya, Paul Pfleiderer, and Stephen A. Ross, 1986, On timing and selectivity, *The Journal of Finance* 41, pp. 715–730.

Agarwal, Vikas, Naveen D Daniel, and Narayan Y Naik, 2011, Do hedge funds manage their reported returns?, *Review of Financial Studies* 24, 3281–3320.

Agarwal, Vikas, and Narayan Y Naik, 2000, Multi-period performance persistence analysis of hedge funds, *Journal of financial and quantitative analysis* 35, 327–342.

Agarwal, Vikas, and Narayan Y Naik, 2004, Risks and portfolio decisions involving hedge funds, *Review of Financial studies* 17, 63–98.

Aragon, George, 2005, Timing multiple markets: theory and evidence from balanced mutual funds, *Unpublished Working Paper, Arizona State University* .

Becker, Connie, Wayne Ferson, David H. Myers, and Michael J. Schill, 1999, Conditional market timing with benchmark investors, *Journal of Financial Economics* 52, pp. 119 – 148.

Bhardwaj, Geetesh, Gary B Gorton, and K Geert Rouwenhorst, 2014, Fooling some of the people all of the time: The inefficient performance and persistence of commodity trading advisors, *Review of Financial Studies* 27, 3099–3132.

Bollen, Nicolas P. B., and Jeffrey A. Busse, 2001, On the timing ability of mutual fund managers, *The Journal of Finance* 56, pp. 1075–1094.

Brain, Carlos W., and Samuel S. Shapiro, 1983, A regression test for exponentiality: Censored and complete samples, *Technometrics* 25, pp. 69–76.

Chance, Don M, and Michael L Hemler, 2001, The performance of professional market timers: daily evidence from executed strategies, *Journal of Financial Economics* 62, pp. 377 – 411.

Chen, Y., 2007, Timing ability in the focus market of hedge funds, *Journal of Investment Management* 5, 66.

Chen, Yong, Wayne Ferson, and Helen Peters, 2010, Measuring the timing ability and performance of bond mutual funds, *Journal of Financial Economics* 98, 72 – 89.

Chen, Z, and PJ Knez, 1996, Portfolio performance measurement: theory and applications, *Review of Financial Studies* 9, 511–555.

Fama, Eugene F, and Kenneth R French, 2010, Luck versus skill in the cross-section of mutual fund returns, *The journal of finance* 65, 1915–1947.

Ferson, Wayne E., and Rudi W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *The Journal of Finance* 51, pp. 425–461.

Fung, Hung-Gay, Xiaoqing Eleanor Xu, and Jot Yau, 2002, Global hedge funds: risk, return, and market timing, *Financial Analysts Journal* 19–30.

Fung, W., and D.A. Hsieh, 1997, Empirical characteristics of dynamic trading strategies: The case of hedge funds, *Review of Financial Studies* 10, 275–302.

Fung, W., and D.A. Hsieh, 2001, The risk in hedge fund strategies: Theory and evidence from trend followers, *Review of Financial Studies* 14, 313–341.

Getmansky, Mila, Andrew W. Lo, and Igor Makarov, 2004, An econometric model of serial correlation and illiquidity in hedge fund returns, *Journal of Financial Economics* 74, 529 – 609.

Goetzmann, William N., Jonathan Ingersoll Jr., and Zoran Ivković, 2000, Monthly measurement of daily timers, *The Journal of Financial and Quantitative Analysis* 35, pp. 257–290.

Harding, Don, and Adrian Pagan, 2002, Dissecting the cycle: a methodological investigation, *Journal of Monetary Economics* 49, 365 – 381.

Henriksson, Roy D., and Robert C. Merton, 1981, On market timing and investment performance. ii. statistical procedures for evaluating forecasting skills, *The Journal of Business* 54, pp. 513–533.

Jensen, Michael C, 1968, The performance of mutual funds in the period 1945–1964, *The Journal of finance* 23, 389–416.

Jensen, Michael C, 1972, Capital markets: Theory and evidence, *The Bell Journal of Economics and Management Science* 357–398.

Jiang, George J., Tong Yao, and Tong Yu, 2007, Do mutual funds time the market? evidence from portfolio holdings, *Journal of Financial Economics* 86, pp. 724 – 758.

Kao, G. Wenchi, Louis T. W. Cheng, and Kam C. Chan, 1998, International mutual fund selectivity and market timing during up and down market conditions, *Financial Review* 33, pp. 127–144.

Kazemi, Hossein, and Ying Li, 2009, Market timing of ctas: An examination of systematic ctas vs. discretionary ctas, *Journal of Futures Markets* 29, 1067–1099.

Lehmann, Bruce N., and David M. Modest, 1987, Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons, *The Journal of Finance* 42, pp. 233–265.

Lunde, A., and A. Timmermann, 2004, Duration dependence in stock prices, *Journal of Business and Economic Statistics* 22, pp. 253–273.

Menkhoff, Lukas, 2010, The use of technical analysis by fund managers: International evidence, *Journal of Banking & Finance* 34, 2573–2586.

Merton, Robert C, 1981, On market timing and investment performance. i. an equilibrium theory of value for market forecasts, *Journal of business* 363–406.

Ohn, Jonathan, Larry W. Taylor, and Adrian Pagan, 2004, Testing for duration dependence in economic cycles, *Econometrics Journal* 7, 528–549.

Pagan, A.R., and K.A. Sossounov, 2003, A simple framework for analysing bull and bear markets, *Journal of Applied Econometrics* 18, pp. 23–46.

Petersen, Mitchell A, 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of financial studies* 22, 435–480.

Scholes, M., and J. Williams, 1977, Estimating betas from nonsynchronous data, *Journal of financial economics* 5, 309–327.

Shapiro, S. S., and M. B. Wilk, 1972, An analysis of variance test for the exponential distribution (complete samples), *Technometrics* 14, pp. 355–370.

Sharpe, William F, 1966, Mutual fund performance, *The Journal of business* 39, 119–138.

Thompson, Samuel B, 2011, Simple formulas for standard errors that cluster by both firm and time, *Journal of Financial Economics* 99, 1–10.

Treynor, J., and K. Mazuy, 1966, Can mutual funds outguess the market?, *Harvard Business Review* 44, pp. 131–136.

Wegscheider, K., 1994, Hierarchische trends, *unpublished* .

Appendix

Threshold Filter by Lunde and Timmermann (2004)

Let I_t be an indicator that takes on the value 1 if the stock market is in a bull state and 0 otherwise. The stock price at the end of period t is x_t. Let λ_1 be a scalar fraction defining the threshold of the movement required to go from a bear to a bull market. Similarly, let λ_2 be the fraction for shifts from a bull market to a bear market. Suppose that at t_0, the stock market is at a local maximum, i.e. $I_{t_0} = 1$. Set $x^{max} = x_{t_0}$ where x_{t_0} is the value of the stock price at time t_0. We can then apply the following filter to classify stock markets:

Step 1: If I_{t-1} at time t equals 1:

1. In the case where $x_t > x^{max}$, the peak is updated so that $x^{max} = x_t$. It is set equal to 1.
2. If $x_t < (1-\lambda_1) \cdot x^{max}$, there is a switch from a bull to a bear market. Retroactively apply $I_t = 0$ since last peak up to time point t.
3. If $x_t > (1 - \lambda_1) \cdot x^{max}$ and $x_t < x^{max}$, it is set equal to 1.

If I_{t-1} at time t equals 0:

1. In the case where $x_t < x^{min}$, the trough is updated so that $x^{min} = x_t$. It is set equal to 0.
2. If $x_t > (1+\lambda_2) \cdot x^{min}$, there is a switch from a bear to a bull market. Retroactively apply $I_t = 1$ since last trough up to time point t.
3. If $x_t < (1 + \lambda_2) \cdot x^{min}$ and $x_t > x^{min}$, it is set equal to 0.

Step 2:

Go back to step 1 until the end of the time series is reached. $\square$

Trend Identification Algorithm by Wegscheider (1994)

Let $T \subseteq \{1, 2, \ldots, N\}$ be a nonlinear subset of observations. The left corner point t_{min} of T being the smallest value, the right cornerpoint t_{max} being the largest value of T. All other points are called inner observations of T. We write $T^{<}$ for $T\{t_{max}\}$. For $t \in T$ with $t > t_{min}$, t_L is the preceding observation of t in T. Formally:

$$t_L = max\{t' : \ t' \in T, \ t' < t\}$$

Similarly, t_R is the subsequent observation of t in T for $t < t_{max}$ $\square$.

Let $(x_t)_{t \in T}$ with $T_0 \subseteq \{1, 2, \ldots, N\}$ be a time series of at least two values. The size of a particular trend is denoted as γ_p and determined as follows:

Step 1: For all $t < t_{max}$ with $x_{t_R} - x_t = 0$, observation t is removed. Let T_1 be the set of remaining observations.

Step 2: If T_1 contains only one element, there is no trend and the iteration is cancelled. All inner points t of T_1 with $x_{t_L} < x_t < x_{t_R}$ and $x_{t_L} > x_t > x_{t_R}$ respectively, share the same trend and are therefore removed. Let T_2 be the set of remaining observations.

Step 3: Let t' be the first observation, at which the smallest difference between two consecutive observations in T_2 starts:

$$t' = min\{t : \ t \in T_2^{<}, \mid x_{t_R} - x_t \mid = min\{\mid x_{s_R} - x_s \mid : \ s \in T_2^{<}\}$$

Trend γ_p is defined by the following arithmetic return:

$$\gamma_p = \frac{x_{t_R'} - x_{t'}}{x_{t'}}$$

When t' and $t_{R'}$ are both inner observations or both corner observations of T_2, the size of the trend is saved and both t' and t'_R are removed.

For $t' = t_{min}$ and $t_R^{'} < t_{max}$, the size of the trend is saved and t' is removed.

For $t_R^{'} = t_{max}$ and $t' > t_{min}$, the size of the trend is saved and t_{max} is removed.

Let T_3 be the set of remaining observations.

Step 4: Set $T_2 = T_3$ and go back to step 3 until T_2 is empty. $\square$

Chapter 5

Adaptive Time Series Momentum: Incorporating Trend Signal Strength and the Performance of Managed Futures[1]

5.1 Introduction

According to BarclayHedge, a public hedge fund database with an extensive coverage of Managed Futures funds, total assets under management (AUM) in the Managed Futures or Commodity Trading Advisors (CTAs) industry stood at 333 billion USD at the end of the third quarter of 2015. This makes the Managed Futures industry the second biggest hedge fund category after Fixed Income Arbitrage.

Until recently, no commonly accepted asset-based benchmarks were available for the CTA industry. Instead, practitioners commonly benchmarked CTAs' performance against manager-based indices. To some extent, the reliance on manager-based benchmarks is related to the challenge with constructing appropriate benchmarks for CTAs, as there is generally

[1]This chapter is based on joint work with Péter Erdős.

no long bias in CTAs' trading strategies. While a manager will generally disclose the markets he or she is active in, the actual position at any point in time will be long or short, depending on the manager's assessment of the prevailing trend in the underlying market.

Benchmarking against peers has its limitations, however. Manager-based benchmarks reflect both the returns to potential market inefficiencies that the constituents in the index attempt to exploit as well as individual managers' skill. Moreover, Fung and Hsieh (2004) point out that hedge fund indices can be expected to inherit some of the biases that are inherent in hedge fund databases. As a consequence, the alpha estimated from such models for any individual manager may not accurately reflect managerial skill.

Instead of benchmarking against peers, an alternative approach that consists of benchmarking managers against a naïve trend-following strategy which is completely asset-based may be more valuable. Moskowitz, Ooi, and Pedersen (2012) are the first to propose a futures-based trading strategy that captures the returns to systematic trend-following in futures markets.[2] The authors coin the observed trend effect time series momentum, and show that time series momentum cannot be explained by the risk factors proposed by Fama and French (1993) and Carhart (1997). Baltas and Kosowski (2013) build on the work of Moskowitz, Ooi, and Pedersen (2012) to suggest a set of the Futures-Based Trend-Following Strategies. Considering weekly and daily strategies in addition to monthly strategies, the authors show that their proposed TSMOM factors significantly improve the explanatory power of multifactor models applied to Managed Futures funds' returns.

In this paper, we contribute to the literature on the performance evaluation of Managed Futures funds in two ways. First, we evaluate the performance of a trend-following strategy that combines short-term time series momentum signals with longer-term time series momentum signals. Whereas a standard time-series momentum approach relying on binary signals does not capture trend strength, aggregating time series momentum signals of different lookback horizons results in a signal that measures the strength of a trend in a particular market. This allows us to allocate to a position in proportion to the signal strength.

We hypothesize that incorporating signal strength may yield a more

[2]For completeness, we note that Fung and Hsieh (2001) propose so-called Primitive Trend-following Factors (PTFS) for trend-following funds. These factors capture the returns to market timing using constructed lookback-straddle prices. To the best of our knowledge, these factors have not gained industrywide traction.

robust time series momentum factor that better anticipates reversals. In addition, incorporating signal strength can be expected to be closer to the actual practices of trend-following asset managers. Coincidentally, the aggregation over a wide range of potential parametrizations avoids an arbitrary choice of certain parametrizations and considerably reduces data mining and calibration concerns. Diversification considerations more generally may be another reason for combining signals over different horizons. Baltas and Kosowski (2013), for example, show that time series strategies over different lookback horizons have low correlations, implying considerable diversification benefits.

To provide some intuition on why diversifying among or combining different trend-following signals may add value and improve the overall performance of a strategy, consider the following hypothetical example. Suppose we have two securities, whose price paths are reported in Table 5.1.

[Table 1 about here.]

We note that both securities have the same *initial* value and *terminal* value, and that the securities' returns exhibit identical levels of volatility over the period considered. In other words, both securities only differ in their interim price path. Application of a simple (long-term) time series momentum strategy over the period t through $t-3$ yields a long signal in both instances. When we include the intermediate signals, however, we observe that the trends in both securities are considerably different. Aggregating all the time series momentum signal suggests that a reversal may be taking place for security A, whereas at t there is a strong and persistent trend in security B. This simple example suggests that aggregating signals over different lookback periods may add value as it captures additional features on the nature of the trend.

Second, in implementing the above approach, we incorporate a number of market frictions and real-life limitations, such as contract-specific transaction costs, the impact of exchange rate risk on contracts' profit-and-loss, and delays between signal generation and trade execution. Earlier work by Hurst, Ooi, and Pedersen (2013) points out the importance of some of these frictions. Incorporating practical implementation issues ensures that the factor is both an investable asset-based factor, which allows a meaningful analysis of stand-alone performance, and that it is investable. The latter implies that the our factor can be used as a somewhat more realistic benchmark for the CTA industry. Not taking these frictions into consideration may raise the bar for managers too much, hampering a meaningful interpretation of fund alpha.

We find that our strategy, which we coin adaptive time series momentum (ATSMOM), matches the stylized facts of manager-based indices along a number of dimensions. Moreover, our benchmark significantly outperforms existing benchmarks in explaining the returns of Managed Futures funds. Decomposing the ATSMOM factor, we find that a second significant factor, which we coin the "speed factor", appears to be present in time series momentum's returns. This speed factor, which we extract using a principal component analysis (PCA) and which buys longer-term and sells shorter-term TSMOM strategies, is similar but not identical to the speed factor proposed by Greyserman and Kaminski (2014). We find that Managed Futures funds tend to load negatively on the speed factor. Interestingly, however, we find that fund alpha is positively related to speed factor exposure.

Turning to performance evaluation using our new factors, we find that smaller Managed Futures funds exhibit a nearly even exposure to every asset class under consideration. At the same time, larger funds tend to overweight more liquid (futures) markets, predominantly Fixed Income. Although our asset-based factors capture much of the return variation of CTA managers, we find that some CTA managers continue to outperform on a risk-adjusted basis.

To investigate the drivers of the observed alpha, we analyze the relationship between risk-adjusted performance and fund characteristics. We find that fund characteristics only account for a small fraction of the cross-sectional variation in fund alphas, suggesting that the estimated alphas are indicative of managerial skill. Interestingly, we also document strong persistence in the estimated alphas, in that good annual performance in one year tends to repeat in the subsequent year. Finally, we find that contemporaneous fund flows do not affect the risk-adjusted performance of managers. This suggests capacity constraints are less of an issue for CTAs. These results echo the findings of Baltas and Kosowski (2013), who rigorously test for capacity constraints in trend-following strategies. Their results suggest that futures markets are liquid enough to accommodate the trading activity of the CTA industry.

5.2 Literature Review

Mutual funds are commonly benchmarked against a combination of market indices and risk factors such as the factors suggested by Fama and French (1993) and Carhart (1997). Similarly, most hedge fund categories are evaluated against Fung and Hsieh's seven-factor (or eight-factor) model (Fung and Hsieh, 2004). While these factor models perform well in explaining the returns of mutual funds and most hedge fund categories, their performance in explaining Managed Futures funds' return variation is limited. Instead, the Managed Futures industry still largely relies on manager-based indices. Such indices reflect the average performance of the selected funds and provide a measure of the industry's performance.[3] This practice is in stark contrast to the above-described practices for mutual funds and other hedge fund categories and may have a number of limitations, as pointed out in the introduction.

There have nevertheless been several studies that attempt to model the returns generated by Managed Futures funds. Fung and Hsieh (2001) are among the first to focus on replicating trend-following hedge funds' returns. The authors suggest the use of primitive trend-following strategies (PTFS) based on lookback straddles, which capture the returns of a market timer. While implementing these factors in practice is possible, Harvey et al. (2016) note that it is neither straightforward nor cheap.

More recently, there has been renewed attention for modelling the returns accruing to Managed Futures funds. Moskowitz, Ooi, and Pedersen (2012) are the first to document, in a systematic manner, the presence of a "trend" effect for a broad range of futures and forward contracts. They coin this effect time series momentum (TSMOM), which relies solely on the continuation of the price direction of the asset under consideration. Moskowitz, Ooi, and Pedersen (2012) show that a portfolio of TSMOM strategies, diversified across different asset classes, consistently delivers large and significant excess returns. Time series momentum is related, but not identical to cross-sectional (or Carhart) momentum which relies on past winners outperforming past losers.

Baltas and Kosowski (2013) build on the work of Moskowitz, Ooi, and Pedersen (2012) to suggest a set of the Futures-Based Trend-Following Strategies. The authors extend the existing literature on time series mo-

[3]For completeness, we note that the Société Générale (formerly Newedge) Trend Indicator index, which relies on a 20/120 moving average crossover futures-based model, is also sometimes used by practitioners to capture the returns accruing to Managed Futures funds.

mentum by considering weekly and daily strategies. Baltas and Kosowski (2013) also provide clear evidence that Managed Futures funds attempt to exploit momentum in the time series domain. In particular, the authors show that their suggested TSMOM factors significantly improve the explanatory power of multifactor models applied to Managed Futures funds' returns and outperform the PTFS suggested by Fung and Hsieh (2001).

Our approach borrows from and extends the works of Moskowitz, Ooi, and Pedersen (2012) and Baltas and Kosowski (2013). In particular, we investigate the economic gains of using more than one or just a few time series momentum signals. The use of multiple signals can be motivated along several lines. First, aggregating a large number of signals results in a combined signal that captures signal strength. This addresses a limitation in existing applications of time series momentum strategies where the binary nature dictates a either a long or short allocation, regardless of the strength of the trend. As a consequence, risk is allocated across different securities and sectors without regard to the strength of the trends in the different markets. It seems reasonable to assume that a reliance on an aggregate or several signals is closer to the CTA industry's practice. Second, an investor ex ante does not know the performance of a particular (e.g. a twelve month) parametrization. From a diversification perspective, it may be more prudent to combine a considerable number of signals.

The choice of the strategy parameters is also an important consideration from a performance evaluation perspective. One can question the investabilitity of a benchmark that is based on just one signal, since such a strategy is underdiversified and suffers from a hindsight bias. This hindsight bias is inherent when relying on specific parametrizations that performed well historically and it may raise the bar for managers too much, as pointed out by Hurst, Ooi, and Pedersen (2012). Combining different candidate signals, however, we avoid having to select a specific set of parameter specifications, thereby reducing model risk while at the same time enhancing 'signal' diversification.

The idea of combining trend signals from different lookback periods matches a recent new avenue in academic research. In particular, Han, Zhou, and Zhu (2016) analyze the economic gains of combining signals from short-, intermediate, and long-term moving average signals in equities. They find that combining the price trend information outperforms the price trends separately. Our work is similar in spirit, but it has a different scope in terms of assets. Additionally, since we focus only on signals of up to one year we do not have to consider price reversals which, literature suggests, tend to occur over horizons beyond one year.

Our adjustment to existing time series momentum strategies implies that our work is also strongly related to other recent contributions that attempt to improve time series momentum strategies. Baltas and Kosowski (2015) investigate the impact of different volatility estimators on the strategy's turnover and find that more efficient estimation of volatility can substantially reduce rebalancing costs. They continue to show that taking into consideration pairwise correlations among assets further improves time series momentum's performance by limiting downside risk.

5.3 Data

We employ data from several distinct data sources. To construct the ATSMOM strategy, we employ daily futures contract data obtained from CSI Data. We retrieve monthly data on Managed Futures funds from the BarclayHedge database. In addition, in-house data from RPM Risk & Portfolio Management AB complements the subsection where we estimate the transaction costs for CTAs.[4]

5.3.1 Futures Data

The futures dataset that we use consists of daily Close Price, Open Interest, and Volume for 98 futures contracts across four asset classes. Individual futures contract data are obtained from CSI Data and cover the period from January 1990 to September 2015. We report the list of futures contracts covered in Table 2. Since some contracts only started trading or were discontinued during the sample period, we also report the period over which each contract is actually included in the subsequent analysis.

[Table 2 about here.]

[Table 2 (cont.) about here.]

Since futures contracts are short-lived contracts that expire at a predetermined date, we first construct a continuous time series of futures prices for each contract. In the online appendix, we describe the particular approach used. The daily returns calculated from the continuous

[4]RPM Risk & Portfolio Management AB, a specialist investment manager based in Stockholm, Sweden is a fund-of-funds specializing in Managed Futures strategies and liquid Global Macro managers.

futures prices, are equivalent to fully collateralized (unleveraged) returns in excess of the risk-free rate (for a thorough discussion, refer to Baltas and Kosowski (2015) and references therein). As such, the daily excess returns are constructed as

$$r_{i,t} = \frac{F_{i,t} - F_{i,t-1}}{F_{i,t}} \tag{5.1}$$

where F_t corresponds to the futures price of asset i at time t. The list of futures contracts that we employ is one of the most comprehensive used in the literature, as we include a number of metal-related futures and a number of currency pairs that are commonly traded by CTAs.

5.3.2 Managed Futures Data

To analyze the relationship between our proposed strategy and Managed Futures funds' performance we collect monthly net-of-fee returns of live and dead funds labeled CTA in the BarclayHedge Database. Although reporting to hedge fund databases is voluntary, Joenväärä, Kosowski, and Tolonen (2012) – in an analysis of the different publicly available hedge fund databases – conclude that BarclayHedge is the most comprehensive hedge fund database, especially for Managed Futures funds. We restrict the data on Managed Futures funds to the period from January 1994 to September 2015. We employ data from January 1994 to mitigate a potential survivorship bias, since most databases only started collecting information on defunct programs from 1994 onwards (see Joenväärä, Kosowski, and Tolonen, 2012).

We filter the sample of funds by looking at their self-declared strategy description and remove funds whose description is not consistent with the definition of CTAs. In the process, we discard duplicates by excluding multiple share classes and focus on the fund's flagship program that attracts the largest assets-under-management (AUM). To account for backfill bias, we drop the first 12 observations (see Kosowski, Naik, and Teo, 2007).[5] We also drop funds with (AUM) below 10 million USD to restrict the set of funds to the investable universe. Finally, we focus on funds that report their returns either in USD or EUR. The EUR-denominated re-

[5] By keeping track of the number of months that are backfilled when a fund is first included in BarclayHedge database, we have tracked backfill bias for the period 2005-2010. For that sample period, the median (average) backfill bias was twelve (fourteen) months.

turns and AUM are converted to USD using the end-of-month EUR/USD spot rate provided by the Bank of England.

We focus on systematic trend-following CTAs, which we select based on funds' self-declared strategy description as well as an analysis of their return characteristics. We focus on systematic trend-following managers as their performance is most clearly related to the concept of time series momentum. These managers do not make discretionary decisions and show a high correlation with manager-based benchmark indices such as the SG Trend Index. These programs are usually diversified and invest across many liquid futures markets. Applying the above adjustments, we obtain a sample of 433 systematic trend-following CTA funds. From this set of funds, we construct both an AUM-weighted and an equal-risk weighted index. Both portfolios are rebalanced monthly.

5.4 Methodology

We construct a portfolio which follows a strategy that we will refer to as ATSMOM, and which is diversified both across time and across asset classes. The aim is to construct a portfolio that is more representative of systematic trend-following CTAs than a time series momentum approach based on a single lookback period. We can imagine that diversification benefits across time and assets result in fund performance that is less sensitive to inevitable trend reversals.

The construction of the ATSMOM builds on the works of Moskowitz, Ooi, and Pedersen (2012) and Baltas and Kosowski (2013) (hereinafter MOP and BK, respectively). Analytically, using daily returns, a diversified TSMOM strategy can be constructed as follows

$$r_{T+1} = \frac{1}{L}\sum_{t=1}^{L} sgn(r_{T-j,T-1,l}) \cdot \frac{0.4/\sqrt{261}}{\sigma_{T-60,T-1,l}} \cdot r_{T+1,l} \tag{5.2}$$

where sgn is the signum function, that is, $sgn(r_{T-j,T-1,l})$ is the sign of the return over the lookback horizon $[T-j, T-j+1, \ldots, T-1]$ lagged two days, L is the number of assets in the strategy and $\sigma_{T-60,T-1,l}$ is the two-day lagged RiskMetrics' standard exponentially weighted moving average (EWMA) estimator of volatility with a 60-day rolling window.[6]

[6]We first convert the daily returns of futures contracts denoted in a foreign currency to USD, since the weighting scheme in Eq. 5.1 is aimed at obtaining a (ex post) level of volatility in USD.

Algebraically, the EWMA estimator in Eq. 5.2 is calculated as follows

$$\sigma_T^2 = (1-\lambda) \cdot \sum_{t=0}^{T} \lambda^{t-1} \cdot (r_t - \bar{r})^2 \tag{5.3}$$

where λ is the decay factor, which we choose such that the center of mass is at around 60 days. We follow MOP in using this simple model for estimating volatility. The correction factor of 0.4 to the estimated volatility in Eq. (2) is suggested by MOP as to achieve an ex ante volatility of 40% per security. This is motivated from the observation that a 40% scaling factor can be expected to yield risk factors with an ex post volatility of around 12% per annum, which roughly matches the volatility of the equity risk factors of Fama and French (1993) (see Moskowitz, Ooi, and Pedersen, 2012).

The ATSMOM strategy is defined as a time series momentum strategy whereby we average the signal for any given security in the portfolio over a wide set of lookback horizons. Algebraically[7],

$$r_{T+1} = \frac{1}{L} \sum_{t=1}^{L} \left(\frac{\sum_{j=10}^{260} sgn(r_{T-j,T-1,l})}{251} \right) \cdot \frac{0.4/\sqrt{261}}{\sigma_{T-60,T-1,l}} \cdot r_{T+1,l} \tag{5.4}$$

We do not consider lookback periods of strictly less than 10 trading days. In the case of such relatively short trading intervals, the high degree of noise makes the type of signal extraction used here unlikely. Momentum trading at such short intervals can be expected to be based on additional information (e.g. order flow) rather than closing prices alone. Such short-term strategies likely also employ intraday rebalancing. Results for a trading strategy that also includes horizons between 1 and 9 days are qualitatively unchanged and are available up on request. This equivalent to a strategy where the strategy trades the net position of every futures contract across the different lookback portfolios.

From eq. (5.4) it is clear that the signal for every futures contract will vary between minus one and plus one (i.e. $S_t \in [-1, 1]$) depending on the strength of the trend. This is a desirable charactistics as a simple TSMOM strategy based on one lookback period can be criticised on the fact that

[7] An alternative way to think about the ATSMOM strategy is by viewing it simply as an equal-weighted portfolio of diversified TSMOM portfolios over different lookback horizons. The overall strategy only trades the net position of every futures contract across the different lookback portfolios.

it is a binary signal. As a result, a standard time-series momentum signal does not capture signal strength. As we illustrated in the introduction, our approach will mechanically allocate more to the futures contracts that exhibit ‘clearer’ trends. When trends start to fade, however, the short-term signals will force the strategy to lower exposure more quickly than in the case of a strategy that only considers one long-term signal, and vice versa. At the portfolio-level, the strategy reduces exposure to markets where trends become less pronounced and adds to futures contracts where trends are or become more pronounced, in a more ‘adaptive’ way than a standard TSMOM strategy based on a single lookback horizon.

In addition to constructing an adaptive TSMOM strategy, we also attempt to improve existing TSMOM strategies or CTA benchmarks along several other dimensions. First, the available benchmarks imply signal generation and trade execution on the same day, that is, for example, signal generation at the close price and entering the market during the same closing session. When rebalancing frequency is low, such as in the case of Moskowitz, Ooi, and Pedersen (2012) who employ monthly rebalancing, the impact of the exact closing price employed may be limited. In our case, however, the impact may be sizeable as we rebalance and thus may shift positions daily.

In line with the work of Hurst, Ooi, and Pedersen (2013), we systematically skip one trading day between signal generation and trade execution. For example, we only enter a position at Tuesday’s closing price if that decision relies on a signal generated based on Monday’s closing price. Similarly, the first day we can close that same position is during Wednesday’s closing session and the return of such a position will be the percentage price difference between Wednesday’s and Tuesday’s closing prices.

Another aspect we consider is the impact of contracts that are traded in a foreign currency, instead of the base currency (USD). We assume that the collateral or margin is always held in the base currency. Thus, *only* the daily profit and loss (P&L) generated from positions in the contracts traded in a foreign currency is exchanged to USD at the daily closing exchange rate. The margin itself, which is held in domestic currency, is not exposed to exchange rate risk (see Appendix A in Koijen et al., 2016). We use the exchange rates provided by the Bank of England or, when these are not available, the exchange rates of the respective central bank to convert the daily P&L of the foreign currency denominated contracts. We can imagine that incorporating trading frictions and exchange rate fluctuations can improve the explanatory power of industry benchmarks.

We should emphasize that the proposed ATSMOM strategy does not trade every diversified TSMOM portfolio (one for each time frame) separately, but rather trades the netted position after aggregating the signals for each constituting TSMOM portfolio. As such ATSMOM is an equally-weighted portfolio of each TSMOM strategy. This way we follow the industry standard and non-negligibly, we substantially reduce the level of transaction costs. The resulting strategy is likely to increase/decrease existing positions only fractionally each day. Only these net changes and the rollover of positions generate transaction costs.[8]

5.5 Results

In this section, we start by estimating the transaction costs for the futures contracts under consideration. Next, we evaluate the performance of the ATSMOM strategy and compare the approach to the futures-based factors suggested by MOP and BK, as well as a number of more traditional risk factors that are used in the context of hedge fund analysis. We also analyze the relationship between our newly constructed factor and systematic trend-following CTAs. Finally, we extend our baseline strategy by decomposing ATSMOM's drivers, which leads to the introduction of a 'speed factor'. We conclude with an analysis of the relationship between our newly proposed factors and CTAs.

5.5.1 Estimation of Transaction Costs

Existing benchmarks, with the exception of the SG Trend Indicator index, do not consider transaction costs incurred executing a systematic trend-following program. To allow for a meaningful performance measurement, we account for transaction costs. A prerequisite to the formation of a CTA benchmark that considers costs is, of course, appropriate estimates of the trading costs typically incurred by CTAs.

[8] When a futures contract is rolled over to a further-dated contract, the strategy closes the nearby contract and opens a position in the new contract. The date of the contract rollover coincides with the rollover used for the construction of the continuous futures (see the Appendix). On such days, turnover is usually much higher than on other days. Daily turnover fairly limited, except in the case of short rate futures. These contracts exhibit very low levels of volatility (0.01% average daily volatility) compared to other contracts (1.2% average daily volatility) and thus a large notional position is needed to obtain the same target level of volatility. Omitting the Eurodollar, the Euribor, and the 90-day bank accepted bill, the turnover equals 29%. Each short rate futures generates an average daily turnover of around 22-23%, whereas, the average turnover for the other contracts is just 0.3%.

To this end, we first estimate the explicit trading costs from actual charges incurred in one of RPM Risk & Portfolio Management's flagship funds over a one-year period from August 2013 through August 2014. Explicit trading costs include gross commissions, clearing fees, exchange fees, NFA (National Futures Association) charges, and brokerage and execution charges. Second, we also need to account for implicit transaction costs arising from the bid-ask spread that traders usually pay market makers for providing liquidity.[9] In line with the standard approach in the literature, in a round trip, the bid-ask spread can be approximated by the tick size. This simplification dates back to Demsetz (1968), who argues that when customers trade through market makers, they will pay the difference between the true price and the bid or ask price on every trade. We therefore employ the reported tick-size for every contract to approximate the implicit transaction cost for every contract.[10]

Ideally, we should re-estimate transaction costs from time to time. Unfortunately, we only have transaction costs data for a very recent period. Following Hurst, Ooi, and Pedersen (2012), we therefore assume that in the first half of the sample period (1991-2002), transaction costs were twice as high as in the second half of the sample period (2003-2015). Table 5.3 reports the estimated transaction costs for each asset class.

[Table 3 about here.]

The results in Table 5.3 clearly illustrate that trading costs vary considerably from asset class to asset class; in basis points of traded notional amount, short-rate futures are the least expensive to trade, though these contracts are also the least volatile. Trading in VIX and grains futures is most expensive. This finding is mainly driven by large tick size indicating

[9] Effective spread estimators (Roll, 1984; Smith and Whaley, 1994) and approaches to estimate bid-ask spread directly from the order book (Locke and Venkatesh, 1997) have also been proposed. Szakmary, Shen, and Sharma (2010) and Locke and Venkatesh (1997) point out, however, that these estimates are close to the tick size. Since estimating the bid-ask spread from the order book is beyond the scope of the current paper, we stick to the simplification that the tick size is a good proxy for the bid-ask spread.

[10] We note that transaction costs are likely to be a nonlinear function of trading volume. In the absence of transaction-level data, however, it is not possible to quantify the relationship. In addition, taking into consideration transaction costs and other frictions such as position limits requires assumptions on the portfolio's size. We refer to the work of Frazzini, Israel, and Moskowitz (2012) for more details on the impact of transaction costs on exploiting asset pricing anomalies. In this study, we assume that transaction costs increase linearly with trading volume.

lower liquidity in these markets. In all markets, except for energy commodities and industrial metals, the half tick size accounts for more than half of the total estimated trading cost. On average, across all markets traded, we find that the bid-ask spread is responsible for almost three quarters of the overall transaction costs.

5.5.2 Adaptive TSMOM's Stand-alone Performance

In Table 5.4 we report performance statistics for the adaptive TSMOM strategy as well as results for the factors suggested by Moskowitz, Ooi, and Pedersen (2012) and Baltas and Kosowski (2013). The diversified time series momentum factor (henceforth MOP) is available from Applied Quantitative Research's (AQR) website. The monthly, weekly, and daily Futures-based Trend-following Benchmarks (FTB, henceforth BK_M, BK_W, and BK_D) are available from Robert Kosowski's website. For the ATSMOM factor, we report the results both gross and net of transaction costs in panel A. The existing benchmarks, in panel B, are gross of transaction costs. All the factors are scaled to 10% volatility for comparison.

[Table 4 about here.]

We observe that the ATSMOM strategy yields somewhat higher minimum and maximum returns than the MOP factor and the BK_D and BK_W proposed by Moskowitz, Ooi, and Pedersen (2012) and Baltas and Kosowski (2013). This suggests that the ATSMOM strategy is successful at limiting downside risk and to allocating more to better performing assets. The lower downside risk is likely to be the consequence of diversification benefits as well as the higher rebalancing frequency. In particular, more frequent rebalancing implies that the strategy will respond more quickly to changes in trends. In contrast, MOP's factor is rebalanced monthly. More frequent rebalancing, however, does not guarantee lower downside risk, as is evident from BK_D's MDD. Taking into account transaction costs, the benefits resulting from the more pro-active nature of the adaptive TSMOM strategy clearly come at a cost. The Sharpe ratio net-of-transaction costs drops to 0.96.

The higher upside of the ATSMOM strategy also translates to a higher skewness and kurtosis. High skewness is consistent with one of the stylized facts of CTAs in that these funds tend to produce positively skewed returns (refer to, among others, Fung and Hsieh (2001), Lamm Jr (2005), and Ding and Shawky (2007)). This feature is also present in the BK_W

and BK_D. Before transaction costs, we find that the ATSMOM strategy reports slightly lower average annual returns than the MOP factor, resulting in a Sharpe ratio of the ATSMOM that does not differ significantly from the Sharpe ratio of MOP (using the approach of Ledoit and Wolf (2008) to test the statistical significance, we obtain a p-value of 0.288).

The focus of this work is, of course, not on stand-alone performance. The results so far simply indicate that our newly proposed benchmark is able to compete with existing benchmarks. Next, we turn to the use of ATSMOM as a benchmark for the Managed Futures industry. Does the adaptive nature of our newly proposed factor better capture Managed Futures funds' performance?

5.5.3 A Benchmark for Managed Futures Funds?

In Panel C of Table 5.4, we report the performance of existing industry indices. These indices are often used by practitioners to benchmark individual managers. The two most commonly used CTA benchmarks are the BarclayHedge CTA Index and the SG (formerly Newedge) CTA Index. BarclayHedge also publishes a large cap index called BTOP 50 and SG a Trend-Following sub-index.[11] In addition to these manager-based indices, SG also constructs an asset-based benchmark called the SG Trend Indicator index which reflects the returns of a strategy that relies on a simple 20/120 moving average crossover model. The index is reported net of transaction costs and a hypothetical 2% management and 20% performance fee.

In addition to the above indices, we also construct an AUM weighted as well as an equal risk-weighted (ERW) index using the systematic trend-following CTAs selected in Section 5.3. Similarly to the other CTA indices, these indices are also far from investable as one cannot rebalance a CTA portfolio on a monthly basis. Lengthy due diligence and legal processes to opening new managers and closing existing managers makes such an approach impractical. Nevertheless, the indices are representative of then-current CTAs. Further, it is reasonable to expect that the TSMOM-based benchmarks are particularly relevant for systematic trend-followers,

[11]The BarclayHedge and SG manager-based indices are equal-weighted. This has the drawback these indices are overweight CTAs that target higher levels of volatility. The manager-based indices are rebalanced once a year. The BarclayHedge CTA index is a broad index of CTAs, some of which are not necessarily trend-followers nor systematic. The SG CTA index includes only the largest 20 CTAs that are open to investment and report performance and AUM on a daily basis.

but may be not for other types of CTAs. While time series momentum benchmarks may also be relevant for discretionary trend-followers, the data set at our disposal only includes 19 discretionary trend-following CTAs that meet the selection criteria. For this reason, we do not include discretionary managers explicitly.

We observe that most of the CTA indices exhibit positive skewness, drawdowns of approximately 15% at 10% annual volatility (with the exception of the BTOP 50), and Sharpe ratios of 0.31 to 0.93. The Trend Indicator strategy reports the highest drawdown, which may be because of the fact that the index employs just one long-term moving average crossover. The industry practice, in contrast, may be rather to apply several different horizons simultaneously, thereby limiting downside risk.

In Figure 1 we plot the 3-year rolling window Sharpe ratio of the different benchmarks reported in Panel A and B of Table 5.4.

[Figure 1 about here.]

The performance of the proposed ATSMOM strategy is almost always somewhere between the slower(-to-react) (MOP TSMOM) and faster (BK Daily) strategies and is less likely to significantly out or underperform the other benchmarks. This is what one would expect from a strategy that allocates both to shorter and longer-term strategies. Longer-term strategies usually outperform shorter-term strategies. This was clearly the case during 2013 through 2015, when the MOP factor clearly outperformed ATSMOM. However, in periods when shorter-term strategies outperform, longer-term strategies tend to suffer. Greyserman and Kaminski (2014) note that it may be difficult, if not impossible, to determine ex ante the horizon that will perform best over a given period. In such an environment, it may be better to trade a wide portfolio of horizons.

To put the performance of the adaptive TSMOM strategy in another perspective, Figure 2 compares the rolling 3-year Sharpe ratio of the adaptive TSMOM strategy, the SG Trend indicator, and the MOP factor, on the one hand, and peer-based indices, on the other hand. We observe that the SG Trend indicator performed better in the early period of the sample, although slightly underperforming the manager-based indices most of the time. The performance of the adaptive strategy follows the performance of trend-following managers more closely, especially in recent years. Both observations are consistent with market participants' sense that the CTA industry is moving towards increased sophistication and diversification.

[Figure 2 about here.]

In Table 5.5, we report the correlation of the different futures-based strategies with the manager indices. With regard to the factors of Baltas and Kosowski (2013), we include a linear combination of the three separate factors which we refer to here as 'Average BK'.

[Table 5 about here.]

Interestingly, the correlation between the adaptive TSMOM strategy and the manager-based indices exceeds the correlation of the average of the BK factors. This suggests that our factor may add value over a combination of the factors of Baltas and Kosowski (2013).

Moving beyond simple summary statistics, we investigate the relationship between our proposed adaptive TSMOM strategy and existing (equity-based) risk factors, the primitive trend-following strategies (PTFS) of Fung and Hsieh (2001) and a number of other recently proposed risk factors in Table 5.6.

[Table 6 about here.]

In column (1) and (2) we report the results for regressions specifications where we regress the monthly (excess) returns of the adaptive TSMOM strategy on the excess returns of the Fama and French (1993) factors and a combination of these and Carhart (1997) cross-sectional momentum factor. We find that the adaptive TSMOM factor produces economically large and significant alphas against existing risk factors, both gross (Panel A) and after transaction costs (panel B). The alphas vary from 9.5% p.a. to up to 13.2% p.a. These results mimic the findings of Moskowitz, Ooi, and Pedersen (2012) that time series momentum is not well explained by existing (equity-based) risk factors.

In column (3), we include the tradable (equity-based) liquidity factor of Pástor and Stambaugh (2003) and find that (equity) liquidity is unrelated to TSMOM. The results in column (4) report the estimates for a regression where we include the PTFS factors of Fung and Hsieh, 2001. In column (5), we report the results for 8-factor model of Fung and Hsieh (2004), where we include all five PTFS factors rather than just the commodities, bonds, and foreign exchange PTFS. While the extended FH model tends to work well for most hedge fund categories (see Fung and Hsieh, 2004), only the PTFS factors are significant in explaining our TSMOM factor. The results corroborate earlier findings that TSMOM is generally unrelated to equity risk factors and that it is only partly related to existing momentum factors such as Carhart's cross-sectional momentum and the lookback straddle based trend-following factors of Fung and Hsieh (2001). In the appendix, we follow the work of Moskowitz, Ooi, and Pedersen (2012), Asness, Moskowitz, and Pedersen (2013), and Koijen et al. (2016) and regress the adaptive TSMOM strategy's returns on a number of macroeconomic, liquidity, volatility, and sentiment variables. We find that variation in these variables does not explain the observed excess returns of the adaptive TSMOM strategy.

Finally, we also regress the strategy's returns against the Global Value and Global (cross-sectional) Momentum factors proposed by Asness, Moskowitz, and Pedersen (2013), which are arguably more appropriate since these factors cover multiple asset classes.[12] We find that both factors perform somewhat better in explaining the variation in our strategy's returns, with both coefficients being significantly positive. The strategy, however, continues to generate a significant and substantial alpha of 5.64% p.a. *vis-à-vis* these factors.

Table 5.7, Panel A, reports the explanatory power of a number of asset-based style regressions, where we regress the most commonly used manager-based CTA indices against commonly used asset-style based hedge fund benchmarks. We consider the period from January 2000 through January 2012, for which data for all variables is available.

[Table 7 about here.]

Consistent with our earlier findings, Fung and Hsieh' PTFS explain up to 30% of the variation in the manager indices. The 10-factor model, which considers other hedge fund asset-based style factors in addition to the PTFS, performs marginally better, although it still only accounts for 20% to 35% of the variation in CTAs' returns. Turning to the SG Trend Indicator, an industry benchmark that has gained some traction among practitioners in the CTA industry, we find that this indicator performs surprisingly well over the sample period considered. Moskowitz, Ooi, and Pedersen (2012) their TSMOM factor also performs consistently across the CTA benchmarks and produces R^2s of around 45%, slightly lower than that of the Trend Indicator. The three-factor model of Baltas and Kosowski (2013) yields comparable results, with adjusted R^2s ranging from 40% to 50%, in line with the authors' findings. The adaptive TSMOM strategy, however, performs better across the board.

[12]We thank an anonymous referee for pointing out this additional analysis.

Next, we also perform 60-month rolling regressions to analyze potential time-variation in the explanatory power of the different asset-based style factors. The explanatory power for the different models vis-à-vis the BarclayHedge index (ERW) is reported in Figure 3.

[Figure 3 about here.]

Two points are worth noting. First, the explanatory power of Fung and Hsieh's 8-factor model that incorporates the all PTFS factors has improved somewhat the last few years, suggesting that CTAs have behaved more like other hedge fund categories in recent years. Two, the ATSMOM factor mimics CTAs' returns more closely in the second half of the sample period.

These results, while tentative, leave unanswered the question of statistical significance. To determine whether the observed increase of our proposed factor in capturing CTAs' returns is meaningful, we compare the adaptive time-series momentum strategy to the model proposed by Baltas and Kosowski (2013).[13] To this end, we first estimate the incremental value added from using the adaptive momentum strategy by calculating the residuals from a regression that regresses the adaptive time series momentum strategy against the Futures-based Trend-following Benchmark (FTB) Strategies. For comparison purposes, we scale all the regressors including the residuals to 10%. We then rerun the specification of Baltas and Kosowski, including the obtained residuals. If the coefficient on the residuals is statistically significant, then this confirms that our proposed factors adds value over and above the FTB. The results are reported in panel B of table 5.7.

Not only do we find that the coefficient is significant at conventional levels, and leads to a meaningful increase in the explanatory power of the models (i.e. a 15 to 20 percentage points increase compared to the initially reported adjusted R-squared, see Panel A), we also observe that the relationship is economically significant. In particular, scaled to the same volatility, we find that the coefficient on the residuals is comparable in magnitude to Baltas and Kosowski's monthly and weekly factor.

[13] We refrain from using an incremental F-test because of potential multicollinearity issues. Table 5.5 indicates that our proposed factor and the average of the FTB exhibit a 0.8 pairwise Pearson correlation.

5.5.4 Decomposing Adaptive Time Series Momentum

Our approach uses TSMOM portfolios with lookback horizons from 10 days to 260 days as the building blocks, with the adaptive TSMOM strategy trading the net position. We can look at these 251 portfolios as separate variables, jointly describing trend-following performance. In an attempt to better understand CTAs' returns, we try to decompose the proposed strategy's returns into its constituent (significant) factors. The question we wish to evaluate here is whether a single factor, which we call the adaptive TSMOM strategy and which is a simple average of the TSMOM strategy portfolios, is enough to fully describe time series momentum strategies in general. The evidence in Greyserman and Kaminski (2014) suggests that there may be other factors beyond ATSMOM driving CTA returns.

One way to address this empirical question is to employ a principal component analysis (PCA) to the constituting TSMOM portfolios. To analyze the statistical significance of the different principal components in time series momentum's returns, we draw 10,000 bootstrapped samples (see Peres-Neto, Jackson, and Somers, 2003) to calculate p-values for the estimated eigenvalues. The eigenvalues are compared to both the broken-stick and Marcenko and Pastur distribution (see e.g., Süss, 2012).

We find that, at the 90% confidence level, both distributions indicate that the first three principal components, corresponding to the three largest eigenvalues, are significant. At the 95% level of significance, the Marcenko-Pastur critical values still point towards three significant components. The broken-stick model, however, suggests that only the first two PCs are significant. Regressing the CTA manager-based indices against the first three PCs, we find that only the first two are significant.

In Figure 4, we plot the loadings of the first two principal components of the 251 horizon portfolios and the corresponding 95% bootstrapped confidence bands applying the bootstrap procedure suggested by Peres-Neto, Jackson, and Somers (2003). The first principal component (PC1) is similar to an equal-weighted portfolio of horizon portfolios, which is consistent with the definition of the adaptive TSMOM strategy. Indeed, the first PC shows a correlation of 0.99 with the strategy's net returns.[14]

[14] ATSMOM, by design, assigns an equal weight to each TSMOM strategy with a lookback window between 10 and 260 days. This implies that there is a significant amount of overlap in the lookback windows. For example, the 10-day window is also part of the 11-day, the 12-day, up to the 260-day window (though it becomes increasingly less important in determining the trend). To generalise, any N-day window is

[Figure 4 about here.]

The second principal component (PC2), however, does not load uniformly on the different constituent portfolios. Instead, Figure 4 indicates that PC 2 is equivalent to a strategy that buys shorter horizon (strategies that react fast to changes in trends) and sells longer horizon momentum strategies (strategies that react slowly to changes in trends). It can therefore be interpreted as a "speed factor", referring to the trading speed of the TSMOM strategies. The factor is close to the opposite of the speed factor in Greyserman and Kaminski (2014), which buys longer (slower) and sells shorter-horizon (faster) momentum strategies. Nevertheless, without loss of generality, we take the negative of PC 2 to get a speed factor similar in Greyserman and Kaminski (2014). Principal components are indifferent to scaling since they are extracted in a way to show zero pairwise correlation.

We know that longer-term momentum strategies outperform their shorter-term counterparts. At the same time, however, longer-term strategies also generate lower skewness (see Table 5.4). The positive average return of the speed factor may thus be a compensation for the lower skewness of longer-term strategies. In that sense, the speed factor can be interpreted as a risk factor. The reasoning that the lower skewness is compensated by the speed factor is related to the arguments provided by Greyserman and Kaminski (2014). They argue that the speed factor is a reward for higher loss tolerance of longer-term momentum strategies.

In Figure 5 we plot the Sharpe ratio for a portfolio that combines ATSMOM with the speed factor, net of transaction costs, as a function of the weight of the speed factor. If the speed factor is scaled to the volatility of the adaptive TSMOM strategy then, through diversification and lower trading costs, the factor contributes positively to the overall performance if its weight is capped at 20% (see Figure 5). Diversification follows from the fact that, by construction, the speed factor has a correlation of zero with the adaptive TSMOM strategy (although the sample correlation may deviate from zero). Thus, calculating the net returns of the speed factor, we assume that it has a (risk) weight equal to 20% of the overall adaptive strategy.

---

part of all longer-term windows. For this reason, the short-term windows are generally more "over-weight" in the overall strategy. This becomes obvious in Figure 4, where PC1 and PC2 load collectively more in the shorter-term signals. These dynamics may explain the large significance of the first two eigenvalues. We thank an anonymous referee for this valuable insight.

[Figure 5 about here.]

We note that, although the ATSMOM strategy is a tradable momentum trading program, PC2 is not yet a tradable factor. This is because PC2 is not net of transaction costs and its composition relies on loadings that are estimated in-sample. Without accounting for real trading conditions, the performance measurement vis-à-vis this factor may be misleading. Therefore, we construct a tradable factor which we henceforth refer to as the speed factor. The weights of the horizon portfolios in the speed factor at any point in time are proportional to the loadings estimated over the entire past history up to the penultimate day, to avoid a look-ahead bias. The initial training period is one year.

It is unlikely that a CTA will trade a strategy similar to the speed factor on a stand-alone basis or separately from a more general TSMOM strategy. It may instead be the case that the speed factor is used as an overlay to complement a more general trend-following strategy, and that only the net positions are traded. From this perspective, only the additional trading costs related to the speed factor need to deducted. In what follows, we discuss the speed factor's performance from this perspective.

To further analyze the newly introduced speed factor, we regress the factor against existing risk factors in Table 5.8.

[Table 8 about here.]

As expected, we find that the speed factor is unrelated to the adaptive TSMOM strategy. At the same time, however, it appears to be related to BK's factors, the PFTS factors, the Carhart cross-sectional momentum factor, and the Stambaugh-Pastor liquidity factor. The positive association with the liquidity factor may be surprising at first sight, especially in light of the earlier finding that the adaptive TSMOM strategy is unrelated to liquidity risk. The speed factor, however, invests in longer-term (slower-to-react) momentum and sells shorter-term (faster-to-react) momentum strategies and can thus be expected to be more exposed to liquidity risk as longer-term systems accommodate slower to a situation when liquidity dries up. On account that the speed factor is an auxiliary factor, we calculate descriptive statistics for the speed factor's and the combined portfolio's returns net of transaction costs in Table 5.9.

[Table 9 about here.]

The speed factor itself underperforms the adaptive TSMOM strategy. Because of its complementary nature, however, stand-alone performance is not that meaningful. We therefore focus on the statistics with regard to the portfolio that allocates 80% to the adaptive TSMOM strategy and 20% to the speed factor. Combining the speed factor with the baseline ATSMOM strategy, we find some improvement for a number of key performance measures compared to the standard adaptive TSMOM strategy. We can conclude that the speed factor adds some value from a portfolio management point of view.

5.5.5 The Speed Factor, Asset Class-based Factors and CTA Performance

With the introduction of the speed factor, we repeat the previous analysis where we regress the various manager-based indices against the newly introduced factors. We also extend the analysis by considering asset-class specific factors for commodity, equity index, fixed income, and foreign exchange futures. The asset class-based factors are scaled to 10% volatility. The results can be found in Table 5.10.

[Table 10 about here.]

As we have already discussed in the previous section, the ATSMOM strategy is able to explain a substantial part of the variation in Managed Futures funds' returns (Table 5.10, Panel A) indicating no abnormal returns among the CTA indices. This suggests that the ATSMOM strategy captures CTAs' trading behavior fairly accurately.

Extending the model with the speed factor increases the fit of most of the regressions, with the exception of the SG indices (Table 5.10, Panel B). The intercepts of the regressions have also increased, but remain statistically insignificant in all but one case. The ERW index generates a significant alpha of 1.69% p.a.

In Panel C of Table 5.10, we report the results for the asset class-based adaptive TSMOM strategies. Applying asset class-based adaptive TSMOM benchmarks has two apparent advantages over a diversified adaptive TSMOM strategy. First, the asset class-based benchmarks improve the explanatory power five percentage points on average. Second, asset class benchmarks allow for a style analysis. Since we have scaled the asset class-based factors to 10% volatility p.a., we can compare the loadings directly. Looking at Table 5.10, Panel C and Panel D, we find that CTAs allocate most to fixed income futures and least to FX and commodity futures. However, the weight of each asset class tends to depend on fund size; large capitalization indices, most of all, the BTOP50 and the AUM-weighted Barclay systematic TF invest more in more liquid markets, i.e., fixed income and less in commodities. Small capitalization managers, gauged by the Barclay CTA and equal risk-weighted Barclay systematic TF indices, invest more evenly across asset classes.

Employing the asset-class based ATSMOM factors and the speed factor, we turn to individual CTAs. In particular, we apply the model to all the individual funds included in the BarclayHedge sample that have at least a one-year track record after inclusion in the database (see Section 5.3). We note that dropping funds that stop reporting before turning two years (one-year of track record in the database in addition to the earlier correction for backfill bias) induces some survivorship bias. Table 5.11 reports the mean and median of the parameter estimates for 335 funds that have produced jointly significant betas at the 10% level of significance.

[Table 11 about here.]

On average, our model is able to explain 40% of the variation in individual CTAs' returns. The average (median) alpha is positive at 0.29% (0.82%) p.a., with 16% of the fund alphas significantly positive and 6% significantly negative. For the funds for which we obtain a significant alpha, we observe considerable variation. Funds with significantly positive alphas generate mean (median) alphas of 4.77% (3.91%) p.a. Funds with significantly negative alphas underperform the adaptive TSMOM strategies by an average (median) of 9.55% (6.56%) p.a.

Interestingly, the Fixed Income adaptive TSMOM factor is significant in 70% of the funds. Thus, CTAs tend to be exposed to fixed income most frequently and this result corroborates with the fact that manager-based indices load most heavily on the Fixed Income factor. The commodity sector is the second most important one, being significant in 64% of the cases. The equity factor is significant in 53%, whereas the FX factor is significant for 48% of the funds. The speed factor is also an important driver of CTA returns being significant in half of the regressions.

Having obtained the alphas versus our proposed factors for the individual CTAs, we continue to investigate the role of fund characteristics in generating alpha. For this particular analysis, we regress the alpha for each fund for every year on yearly fund characteristics that include lagged alpha, fund size, fund age, a standard measure of fund flow, R^2 and the relative factor exposures of the performance regressions, the level of management and incentive fees, margin-to-equity (ME) ratio, and round turns per million dollars per year.

The alphas are estimated and therefore subject to measurement error. If we do not correct for this, the measurement error will generate heteroscedasticity in the panel regression residuals and it may cause standard significance tests to be invalid. To correct for potential heteroscedasticity, we weight each observation by the reciprocal of the standard errors of the performance regressions, as in Dahlquist, Engström, and Söderlind (2000).

Table 5.12 reports the results controlling for time-fixed effects. In column (1) we omit the margin-to-equity (ME) ratio and round turns per million USD per year statistics, as they are only available for a subset of CTAs. We run a specification that includes the ME ratio and round turns per million in column (2) and (3) of Table 5.12 , respectively.

[Table 12 about here.]

The results in Table 5.12 suggest strong momentum in Managed Futures funds' performance. CTAs that outperformed our benchmark portfolios in the previous year tend to repeat that superior performance the following year. Fund size appears to negatively affect risk-adjusted performance. Somewhat surprising though, aging is positively related to better alphas. However, for instance, the expected risk-adjusted performance of a five year old CTA that has 1 billion USD under management is, ceteris paribus, 1.7% p.a. less than that of a CTA that manages only 10 million USD but it is only two years old indicating that interpreting one of the variables alone can be misleading.

Contemporaneous fund flows do not affect risk-adjusted performance. This suggests that capacity constraints are less an issue for CTAs. Adding the R^2s of the performance regressions, we test and reject the hypothesis in Sun, Wang, and Zheng (2012) that hedge funds whose returns are less explainable by risk factors bear more managerial skills. In contrast, funds that engage in pure trend-following approaches tend to generate higher risk-adjusted performance. Thus, alpha does not appear to derive from being less mainstream, but from other sources. This may include superior risk management, better trade execution, and lower explicit transaction costs.

The factor weights are simply calculated from absolute loadings in the individual performance regressions. All else equal, we find that higher equity momentum exposure is likely to result in higher risk-adjusted performance. In contrast, funds with higher allocations to Fixed Income TF strategies tend to generate lower alpha. Interestingly, CTAs that have higher exposure to the speed factor significantly outperform those who have less exposure. The speed factor exposure is likely to be a proxy for the level of sophistication of the manager, since our results suggest that there is some benefit from allocating to the speed factor in terms of diversification and lower transaction costs. All in all, asset exposure, i.e., style is partly accountable for superior risk-adjusted performance.

Higher margin usage over capital invested (ME ratio) appears to be a sign of better performance, most probably through economies of scale. This result suggests that higher risk-taking does not, per se, imply inferior risk management and thus poorer performance. Finally, more trading in terms of rounds per million USD per year does not affect risk-adjusted performance.

Only a small part of the cross-sectional variation in estimated alphas is attributable to fund characteristics such as past performance, fund age, fund size, fees, and style. We conclude that the alphas obtained vis-à-vis our new risk factors can, to some extent, be interpreted as capturing managerial skill.

5.6 Conclusion

In this paper we propose a time series momentum strategy that changes the exposure to futures markets more dynamically by aggregating time series momentum signals over a wide range of horizons. This way, the model increases the allocation to the markets where trends are more well-behaved and decreases exposure to the markets where trends are reversing. We find that our approach better explains Managed Futures funds' reported returns. As such, our approach can aid practitioners in benchmarking and manager selection. We also find that a subset of funds continues to exhibit positive alpha vis-à-vis our new risk factors. Moreover, the abnormal returns of these funds can only be partly explained by observable fund characteristics and thus appear indicative of skill.

Importantly, we document strong momentum in CTA risk-adjusted performance, as stellar performance in one year tends to repeat in the subsequent year, and find evidence that fund size is negatively, whereas fund age is positively related to risk-adjusted performance. Fund style, i.e., asset class exposure and the applied trading strategy, also contributes to CTA alphas. Contemporaneous fund flows, in contrast, do not affect risk-adjusted performance, suggesting capacity constraints are less an issue for CTAs. Higher management and performance fees do not signal prospect for better performance.

Bibliography

Asness, Clifford S., Tobias J. Moskowitz, and Lasse Heje Pedersen, 2013, Value and momentum everywhere, *The Journal of Finance* 68, 929–985.

Baltas, Nick, and Robert Kosowski, 2013, Momentum strategies in futures markets and trend-following funds .

Baltas, Nick, and Robert Kosowski, 2015, Demystifying time-series momentum strategies: Volatility estimators, trading rules and pairwise correlations .

Carhart, Mark M., 1997, On persistence in mutual fund performance, *The Journal of Finance* 52, 57–82.

Dahlquist, Magnus, Stefan Engström, and Paul Söderlind, 2000, Performance and characteristics of swedish mutual funds, *Journal of Financial and Quantitative Analysis* 35, 409–423.

Demsetz, Harold, 1968, The cost of transacting, *The Quarterly Journal of Economics* 82, 33–53.

Ding, Bill, and Hany A. Shawky, 2007, The performance of hedge fund strategies and the asymmetry of return distributions, *European Financial Management* 13, 309–331.

Elaut, Gert, Michael Frömmel, and John Sjödin, 2015, Crystallization: A hidden dimension of cta fees, *Financial Analysts Journal* 71, 51–62.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3 – 56.

Frazzini, Andrea, Ronen Israel, and Tobias J Moskowitz, 2012, Trading costs of asset pricing anomalies .

Fung, William, and David A. Hsieh, 2001, The risk in hedge fund strategies: Theory and evidence from trend followers, *Review of Financial Studies* 14, 313–341.

Fung, William, and David A Hsieh, 2004, Hedge fund benchmarks: A risk-based approach, *Financial Analysts Journal* 60, 65–80.

Greyserman, Alex, and Kathryn Kaminski, 2014, *Trend Following with Managed Futures: The Search for Crisis Alpha* (John Wiley & Sons).

Han, Yufeng, Guofu Zhou, and Yingzi Zhu, 2016, A trend factor: Any economic gains from using information over investment horizons?, *Journal of Financial Economics* 122, 352 – 375.

Harvey, Campbell R, Sandy Rattray, Andrew C Sinclair, and Otto Van Hemert, 2016, Man vs. machine: Comparing discretionary and systematic hedge fund performance .

Hurst, Brian, Yao Hua Ooi, and Lasse H Pedersen, 2012, A century of evidence on trend-following investing, *AQR white paper* .

Hurst, Brian, Yao Hua Ooi, and Lasse Heje Pedersen, 2013, Demystifying managed futures, *Journal of Investment Management* 11, 42–58.

Joenväärä, Juha, Robert Kosowski, and Pekka Tolonen, 2012, Revisiting stylized facts about hedge funds, *Imperial College Business School* .

Koijen, RSJ, TJ Moskowitz, LH Pedersen, and EB Vrugt, 2016, Carry. fama-miller working paper, *Available at SSRN 2298565* .

Kosowski, Robert, Narayan Y. Naik, and Melvyn Teo, 2007, Do hedge funds deliver alpha? a bayesian and bootstrap analysis, *Journal of Financial Economics* 84, 229 – 264.

Lamm Jr, R McFall, 2005, The answer to your dreams? investment implications of positive asymmetry in cta returns, *The Journal of Alternative Investments* 7, 22–32.

Ledoit, Oliver, and Michael Wolf, 2008, Robust performance hypothesis testing with the sharpe ratio, *Journal of Empirical Finance* 15, 850 – 859.

Locke, Peter R., and P. C. Venkatesh, 1997, Futures market transaction costs, *Journal of Futures Markets* 17, 229–245.

Moskowitz, Tobias J., Yao Hua Ooi, and Lasse Heje Pedersen, 2012, Time series momentum, *Journal of Financial Economics* 104, 228 – 250, Special Issue on Investor Sentiment.

Pástor, Luboš, and Robert F Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

Peres-Neto, Pedro R., Donald A. Jackson, and Keith M. Somers, 2003, Giving meaningful interpretation to ordination axes: Assessing loading significance in principal component analysis, *Ecology* 84, 2347–2363.

Roll, Richard, 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *The Journal of Finance* 39, 1127–1139.

Smith, Tom, and Robert E. Whaley, 1994, Estimating the effective bid/ask spread from time and sales data, *Journal of Futures Markets* 14, 437–455.

Sun, Zheng, Ashley Wang, and Lu Zheng, 2012, The road less traveled: Strategy distinctiveness and hedge fund performance, *Review of Financial Studies* 25, 96–143.

Süss, Stephan, 2012, The pricing of idiosyncratic risk: evidence from the implied volatility distribution, *Financial Markets and Portfolio Management* 26, 247–267.

Szakmary, Andrew C., Qian Shen, and Subhash C. Sharma, 2010, Trend-following trading strategies in commodity futures: A re-examination, *Journal of Banking & Finance* 34, 409 – 426.

Table 5.1: Example Aggregating Trend-following Signals

	Security A	Security B	$r_{t,A}$	$r_{t,B}$
$t-3$	90	90		
$t-2$	130	83	44.44%	-8.28%
$t-1$	140	120	7.69%	45.37%
t	125	125	-10.71%	4.17%
	Signal	Signal	σ_A	σ_B
$Sign(t-1,t)$	-1	1	28.08%	28.08%
$Sign(t-2,t)$	-1	1		
$Sign(t-3,t)$	1	1		

Table 5.2: Summary Statistics Futures

	Start	End	Cur	$\mu(\%)$	$\sigma(\%)$	Skew.	Kurt.	MDD (%)	SR
Mexican Peso	04-95	09-15	USD	4.82	10.52	-0.97	6.73	-31.20	0.45
Swiss Franc	01-90	09-15	USD	1.11	11.24	0.10	3.92	-49.41	0.10
British Pound	01-90	09-15	USD	1.49	9.19	-0.60	5.41	-29.41	0.16
Canadian Dollar	01-90	09-15	USD	0.40	7.87	-0.34	6.28	-28.21	0.05
Japanese Yen	01-90	09-15	USD	-1.01	10.89	0.60	6.03	-61.65	-0.09
Australian Dollar	01-90	09-15	USD	2.57	11.52	-0.33	4.75	-41.30	0.22
US Dollar Index	01-90	09-15	USD	-0.71	8.53	0.36	3.78	-44.33	-0.08
Euro FX	05-98	09-15	USD	0.27	10.37	-0.03	3.82	-32.35	0.03
SA Rand	05-97	09-15	USD	1.82	16.42	-0.18	3.73	-46.89	0.11
Brazilian Real	11-95	09-15	USD	4.44	18.42	-1.47	13.38	-53.45	0.24
USD/SEK	05-00	09-15	SEK	0.06	11.85	0.17	3.42	-45.53	0.01
USD/NOK	05-00	09-15	NOK	-0.86	11.78	0.49	4.30	-50.80	-0.07
NZ Dollar	05-97	09-15	USD	3.12	13.36	-0.16	4.30	-41.34	0.23
AUD/NZD	05-99	09-15	NZD	-0.74	7.74	0.02	2.79	-28.88	-0.10
AUD/Japan Yen	05-02	09-15	JPY	7.35	15.11	-0.88	6.24	-42.61	0.47
Euro FX/ Yen	01-99	09-15	JPY	2.88	12.49	-0.55	5.12	-40.58	0.23
EUR/Nok	09-11	09-15	NOK	3.55	6.80	0.11	2.46	-8.55	0.52
EUR/SEK	06-11	09-15	SEK	-0.31	4.68	-0.81	4.05	-11.34	-0.07
EUR/GBP	01-99	09-15	GBP	-0.34	8.16	1.63	12.32	-27.73	-0.04
EUR/CHF	01-99	09-15	CHF	-0.88	6.58	-2.60	19.03	-34.66	-0.14
CAC-40 Index	01-90	09-15	EUR	4.59	19.29	-0.34	3.20	-62.89	0.23
Nikkei 225	09-90	09-15	USD	0.42	21.60	-0.11	3.37	-77.47	0.02
Russell 2000	02-93	09-15	USD	7.88	19.12	-0.49	4.19	-53.95	0.40
S&P Midcap 400	02-92	09-15	USD	9.40	16.61	-0.66	5.30	-52.79	0.54
Hang Seng	01-90	09-15	HKD	12.24	26.13	0.25	5.34	-58.90	0.44
DAX	11-90	09-15	EUR	7.12	20.74	-0.51	4.88	-71.72	0.33
S&P 500	01-90	09-15	USD	6.67	14.57	-0.62	4.26	-58.65	0.45
Topix Index	04-90	09-15	JPY	0.21	19.93	-0.17	4.07	-73.13	0.01
FTSE 100 Index	01-90	09-15	GBP	3.44	14.52	-0.40	3.45	-52.82	0.23
Swiss Market	11-90	09-15	CHF	9.15	15.80	-0.59	4.44	-52.65	0.56
Ibex 35 Index	04-92	09-15	EUR	7.87	21.77	-0.22	3.62	-59.23	0.35
MIB 30 Stock	11-94	09-15	EUR	4.61	22.50	0.15	3.66	-68.88	0.20
Nasdaq 100	04-96	09-15	USD	11.18	26.73	-0.27	4.09	-83.03	0.40
MSCI Taiwan	01-97	09-15	USD	5.54	26.56	0.13	3.85	-64.71	0.20
DJ Industrial Avg	10-97	09-15	USD	5.01	14.88	-0.63	4.31	-49.75	0.33
KOSPI 200 Index	01-98	09-15	KRW	10.43	28.99	0.43	4.07	-58.55	0.34
DoJStoxx 50	06-98	09-15	EUR	1.45	16.66	-0.52	3.82	-66.68	0.09
DJ Euro Stoxx	06-98	09-15	EUR	2.40	19.68	-0.43	3.80	-64.00	0.12
S&P Canada 60	09-99	09-15	CAD	5.24	14.93	-0.71	4.66	-51.85	0.34
CBOE VIX	03-04	09-15	USD	-30.93	62.44	1.95	9.15	-99.89	-0.59
OMX	10-92	09-15	SEK	11.87	21.71	0.04	4.72	-72.40	0.52
US MSCI EAFE	09-10	09-15	USD	4.81	15.64	-0.36	3.16	-24.49	0.30
Amsterdam EOE	10-92	09-15	EUR	7.70	19.62	-0.74	4.81	-68.87	0.38
NYSE Comp	01-90	09-11	USD	5.04	14.70	-0.81	5.11	-57.40	0.34
All Ordinary SPI	01-90	09-01	AUD	2.93	14.26	-0.31	2.83	-28.56	0.20
SPI 200	05-00	09-15	AUD	3.72	13.25	-0.75	3.66	-51.85	0.28

Table 5.2: Summary Statistics Futures (*Cont.*)

	Start	End	Cur	$\mu(\%)$	$\sigma(\%)$	Skew.	Kurt.	MDD (%)	SR
Treasury Bonds	01-90	09-15	USD	5.46	9.31	0.10	5.16	-15.83	0.57
Canada 10Y Gov	01-90	09-15	CAD	4.31	6.00	-0.03	3.30	-14.80	0.71
3M-Eurodollar	01-90	09-15	USD	0.19	0.22	0.68	5.77	-0.68	0.87
10-YR Treasury	01-90	09-15	USD	4.47	5.94	0.13	4.73	-11.69	0.74
Japan 10Y Gov	04-90	09-15	JPY	3.76	4.18	-0.58	7.04	-9.59	0.89
Long Gilt	01-90	09-15	GBP	3.64	6.81	0.00	3.44	-15.65	0.53
US 2-YR Treasury	06-90	09-15	USD	1.60	1.64	0.26	3.56	-3.82	0.97
US 10 YR Bonds	01-90	09-15	AUD	4.64	7.87	-0.01	3.26	-23.59	0.58
US 90-Day Bill	01-90	09-15	AUD	0.13	0.25	0.43	6.61	-0.61	0.52
US 3 Year Bonds	01-90	09-15	AUD	2.19	3.41	-0.01	5.05	-8.75	0.64
US 5-YR Treasury	01-90	09-15	USD	3.23	4.02	0.10	3.91	-8.52	0.79
Muni Note Index	01-90	03-06	USD	5.23	6.79	-0.52	3.92	-16.66	0.76
Euro Buxl	10-98	09-15	EUR	6.19	10.73	0.76	5.05	-17.15	0.56
German Bund	10-98	09-15	EUR	4.06	5.27	0.11	2.85	-9.93	0.76
German Bobl	10-98	09-15	EUR	2.60	3.17	-0.02	2.75	-7.42	0.81
German Schatz	10-98	09-15	EUR	0.90	1.31	0.16	4.00	-4.01	0.69
3Y Korean Bond	09-99	09-15	KRW	2.72	3.11	0.39	5.30	-4.86	0.86
PIBOR	01-90	06-99	EUR	-0.01	0.35	-1.47	10.81	-1.56	-0.02
3M Euribor	09-98	09-15	EUR	0.08	0.15	2.38	21.61	-0.52	0.53
Gas Oil	01-90	09-15	USD	11.13	32.05	0.48	5.15	-73.39	0.33
Nat Gas	04-90	09-15	USD	-11.63	48.05	0.57	4.63	-99.81	-0.26
Brent Crude	01-90	09-15	USD	12.10	33.39	0.60	6.76	-75.63	0.34
Heating Oil	06-06	09-15	USD	-4.13	28.23	-0.19	3.94	-70.00	-0.15
Light Crude	01-90	09-15	USD	5.94	33.65	0.44	5.26	-87.15	0.17
Unleaded Gas	01-90	12-06	USD	18.05	36.80	0.84	5.93	-63.18	0.46
Rbob Electronic	10-05	09-15	USD	4.73	33.08	-0.56	5.60	-70.44	0.14
Copper	01-90	09-15	USD	8.19	25.72	-0.03	5.71	-63.90	0.31
Platinum	01-90	09-15	USD	4.40	20.23	-0.55	6.52	-62.28	0.21
Silver	01-90	09-15	USD	4.44	28.45	0.12	3.87	-71.55	0.15
Gold	01-90	09-15	USD	2.05	15.77	0.18	4.25	-61.55	0.13
Palladium	01-90	09-15	USD	10.94	32.68	0.47	6.68	-86.15	0.32
Live Cattle	01-90	09-15	USD	0.43	13.15	-0.69	5.81	-45.11	0.03
Live Hogs	01-90	09-15	USD	-5.02	24.49	-0.08	3.63	-94.06	-0.21
Pork Bellies	01-90	07-11	USD	6.58	38.09	0.84	4.61	-80.00	0.17
Feeder Cattle	01-90	09-15	USD	3.16	13.59	-0.47	5.24	-38.61	0.23
Corn	01-90	09-15	USD	-2.08	26.16	0.32	3.96	-84.50	-0.08
Oat	01-90	09-15	USD	-0.09	29.49	0.65	4.66	-88.85	0.00
Soybeans	01-90	09-15	USD	5.67	23.49	-0.01	3.68	-50.50	0.24
Soybean Meal	01-90	09-15	USD	12.46	25.81	0.46	4.24	-43.72	0.46
Soybean Oil	01-90	09-15	USD	-0.48	24.34	0.13	4.64	-72.25	-0.02
Wheat W	01-90	09-15	USD	-4.88	27.69	0.46	4.81	-94.44	-0.18
Wheat	01-90	09-15	USD	0.03	27.21	0.51	4.65	-82.15	0.00
Cocoa	01-90	09-15	USD	0.34	29.02	0.49	4.17	-90.23	0.01
Cotton No. 2	01-90	09-15	USD	-1.83	26.19	0.26	3.87	-93.14	-0.07
Coffee	01-90	09-15	USD	-1.28	37.88	1.21	6.19	-94.21	-0.03
Orange Juice	01-90	09-15	USD	-3.98	30.04	0.48	4.35	-91.99	-0.14
Sugar No. 11	01-90	09-15	USD	2.59	30.71	0.26	3.59	-72.49	0.08
Lumber	01-90	09-15	USD	-6.22	31.10	0.45	4.16	-97.52	-0.21
Nickel	01-90	09-15	USD	7.79	33.17	0.24	3.52	-79.39	0.23
Aluminum	10-92	09-15	USD	3.55	18.46	-0.34	7.23	-60.47	0.19
Lead	01-90	09-15	USD	7.23	26.52	-0.01	4.34	-72.70	0.26
Zinc	01-90	09-15	USD	4.21	24.56	-0.03	4.84	-74.94	0.17

Table 5.3: Estimated Transaction Costs In Basis Points Of Notional Value

		Explicit Trading Costs		Half Tick Size		One-Way Transaction Costs	
		1991-2002	2003-2015	1991-2002	2003-2015	1991-2002	2003-2015
1	Developed Currencies	0.77	0.38	0.98	0.49	1.75	0.87
2	Emerging Currencies	0.30	0.15	1.89	0.94	2.18	1.09
3	Developed Equities	0.65	0.33	1.22	0.61	1.87	0.94
4	Emerging Equities	0.42	0.21	2.37	1.18	2.79	1.39
5	Bonds	0.20	0.10	1.45	0.72	1.65	0.83
6	Short Rates	0.03	0.01	0.14	0.07	0.17	0.08
7	Energy Commodities	1.13	0.56	1.11	0.56	2.24	1.12
8	Industrial Metals	1.30	0.65	0.90	0.45	2.19	1.10
9	Precious Metals	0.55	0.27	1.00	0.50	1.55	0.78
10	Grains	0.39	0.19	5.23	2.62	5.62	2.81
11	Livestock	0.62	0.31	2.72	1.36	3.34	1.67
12	Softs	0.40	0.20	2.55	1.27	2.95	1.48
13	VIX	-	0.80	-	2.72	-	3.52

This table reports the estimated transaction cost for the different types of futures contracts in the dataset. The first two columns report explicit trading costs. Column 3 and 4 report implicit transaction costs, which we proxy using the reported tick size. The final two columns indicate the one-way transaction cost assumed. Following Hurst, Ooi, and Pedersen (2012) transaction costs during the first half of the sample period are assumed to be twice as high as during the second half of the sample period.

Table 5.4: Descriptive Statistics for Market and Manager-based CTA Benchmarks

	Panel A - ATSMOM		Panel B - Existing Asset-Based Benchmarks				
	ATSMOM - Gross	ATSMOM - net of TCs	MOP	BK_M	BK_W	BK_D	BK Avg
Min (%)	-5.78	-5.91	-8.04	-7.11	-8.43	-7.81	-5.70
Max (%)	15.26	15.22	10.50	8.88	15.29	12.38	14.95
Ann. Mean (%)	11.05	9.62	12.42	12.67	10.98	9.00	14.28
Ann. Median (%)	9.05	7.73	12.61	13.80	10.16	6.62	8.55
Ann. Std (%)	10.00	10.00	10.00	10.00	10.00	10.00	10.00
Skewness	0.70	0.71	0.04	-0.20	0.63	0.77	0.90
Kurtosis	5.55	5.61	3.26	3.12	5.85	4.81	5.56
Sharpe	1.11	0.96	1.24	1.27	1.10	0.90	1.43
Sortino	2.26	1.93	2.24	2.07	2.09	1.95	3.07
Max DD (%)	-11.24	-14.13	-13.23	-14.96	-10.17	-13.12	-12.69
Period	01/92-09/15	01/92-09/15	01/92-09/15	01/92-01/12	01/92-01/12	01/92-01/12	01/'92-01/'12

	Panel C - Existing Industry Benchmarks						
	Barclay CTA	BTOP 50	SG CTA	SG Trend	SG Trend Indicator	BH Syst. TF AUM	BH Syst. TF ERW
Min (%)	-6.55	-8.30	-9.03	-9.53	-6.28	-7.13	-6.13
Max (%)	8.86	10.86	9.66	9.62	11.09	10.41	9.94
Ann. Mean (%)	6.74	3.99	4.36	3.99	3.13	7.07	9.25
Ann. Median (%)	3.10	1.26	2.84	5.34	2.71	6.46	7.13
Ann. Std (%)	10.00	10.00	10.00	10.00	10.00	10.00	10.00
Skewness	0.37	0.34	0.13	0.13	0.55	0.33	0.29
Kurtosis	3.32	3.71	3.34	3.65	3.87	3.28	3.01
Sharpe	0.67	0.40	0.44	0.40	0.31	0.71	0.93
Sortino	1.31	0.74	0.80	0.72	0.62	1.42	1.88
Max DD (%)	-13.67	-21.62	-14.70	-14.77	-28.94	-11.61	-14.06
Period	01/92-09/15	01/92-09/15	01/00-09/15	01/00-09/15	01/00-09/15	01/94-09/15	01/94-09/15

This table reports summary statistics for gross and net of transaction costs of the adaptive TSMOM strategy in Panel A. Panel B reports the summary statistics for the TSMOM factor of Moskowitz, Ooi, and Pedersen (2012) and the factors proposed by Baltas and Kosowski (2013). Panel C reports the summary statistics for a number of industry benchmarks. The BarclayHedge indices are based on the sample of systematic trend followers selected in Section 5.3. All time series in the table have been adjusted to 10% annualized volatility for comparison purposes.

Table 5.5: Correlation between industry benchmarks and market based indices

	ATSMOM	MOP	Average BK	SG Trend	Barclay	BTOP50	SG CTA	SG Trend	BH Sys. TF (AUM)	BH Sys. TF (ERW)
ATSMOM	1.00									
MOP	0.77	1.00								
Average BK	0.80	0.65	1.00							
SG Trend	0.78	0.60	0.59	1.00						
Barclay	0.77	0.66	0.62	0.70	1.00					
BTOP50	0.77	0.65	0.60	0.69	0.93	1.00				
SG CTA	0.78	0.65	0.59	0.73	0.93	0.98	1.00			
SG Trend	0.78	0.68	0.60	0.72	0.93	0.97	0.97	1.00		
BH Sys. TF (AUM)	0.80	0.70	0.67	0.71	0.94	0.96	0.95	0.97	1.00	
BH Sys. TF (ERW)	0.82	0.68	0.68	0.74	0.97	0.92	0.93	0.93	0.96	1.00

This table reports the correlation between the market-based indices and the industry benchmarks.

Table 5.6: Adaptive TSMOM strategy against existing risk factors

	Panel A: Gross of Transaction Costs					
	(1)	(2)	(3)	(4)	(5)	(6)
MKT	-0.1344*	-0.0605	-0.0565		-0.0691	
	(0.070)	(0.069)	(0.066)		(0.063)	
SMB	0.0218	-0.0083	-0.0045		0.0589	
	(0.054)	(0.055)	(0.055)		(0.048)	
HML	-0.0228	0.0473	0.0350			
	(0.070)	(0.067)	(0.070)			
MOM		0.1943***	0.1964***			
		(0.042)	(0.043)			
Liquidity factor			-0.0865			
			(0.056)			
PTFSBD				0.0092	0.0010	
				(0.015)	(0.014)	
PTFSFX				0.0334**	0.0288**	
				(0.014)	(0.013)	
PTFSCOM				0.0455***	0.0441***	
				(0.015)	(0.015)	
PTFSIR				0.0007	0.0008	
				(0.012)	(0.012)	
PTFSSTK				0.0427**	0.0408**	
				(0.018)	(0.018)	
EM					0.0495	
					(0.043)	
Bond Factor					-0.0157	
					(0.012)	
Credit Spread					1.2142	
					(1.460)	
Global VAL						0.3489**
						(0.137)
Global MOM						0.8587***
						(0.123)
Constant	0.0104***	0.0089***	0.0094***	0.0121***	0.0118***	0.0060***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
R-squared	0.033	0.118	0.129	0.195	0.237	0.239

	Panel B: Net of Transaction Costs					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.0091***	0.0076***	0.0081***	0.0108***	0.0105***	0.0047***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
R-squared	0.032	0.114	0.125	0.193	0.236	0.236

Adaptive TSMOM strategy's returns are regressed against existing risk factors. Panel A reports the results for strategy gross of transaction costs. Panel B reports the alpha for the same regressions using the net-of-transaction costs strategy returns. The risk factors for the Fama and French (1993) and Carhart (1997) models have been downloaded from Kenneth French's website. The Pástor and Stambaugh (2003) traded liquidity factor has been obtained from Lubos Pastor's website. The Fung and Hsieh (2001) factors have been taken from David A. Hsieh's Hedge Fund Data Library. The Global value and Global Momentum Factor have been taken from AQR's website. *, **, and *** denote significance at the 90, 95, and 99% level, respectively. Robust standard errors in parentheses.

Table 5.7: Explanatory power market-based indices

Panel A: Adjusted R-squareds

	Barclay	BTOP50	SG CTA	SG Trend	BH Syst TF (AUM)	BH Syst TF (ERW)
PTFS	0.294	0.190	0.199	0.153	0.218	0.303
F&H 10-Factor	0.329	0.251	0.246	0.198	0.266	0.344
Trend Indicator	0.492	0.482	0.540	0.522	0.507	0.549
MOP	0.431	0.430	0.421	0.459	0.484	0.458
BK	0.413	0.399	0.396	0.419	0.489	0.490
ATSMOM	0.592	0.595	0.604	0.612	0.649	0.658

Panel B: Comparison Explanatory Power BK and ATSMOM

	Barclay	BTOP50	SG CTA	SG Trend	BH Syst TF (AUM)	BH Syst TF (ERW)
BK_M	0.2692***	0.4169***	0.4028***	0.4680***	0.5580***	0.2717***
	(0.045)	(0.053)	(0.052)	(0.055)	(0.057)	(0.039)
BK_W	0.2025***	0.2840***	0.3033***	0.3015***	0.3340***	0.2090***
	(0.061)	(0.076)	(0.077)	(0.082)	(0.091)	(0.051)
BK_D	0.0895*	0.1009	0.0837	0.0723	0.1979***	0.1055***
	(0.049)	(0.063)	(0.068)	(0.070)	(0.073)	(0.038)
Resid	0.2966***	0.4584***	0.4756***	0.4771***	0.4969***	0.2866***
	(0.052)	(0.079)	(0.075)	(0.083)	(0.092)	(0.044)
Constant	-0.0006	-0.0040**	-0.0037**	-0.0045**	-0.0023	-0.0003
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
adj. R^2	0.593	0.596	0.607	0.613	0.652	0.676

Panel A reports the adjusted-R^2s for regressions where we regress CTA manager-based indices against the asset-style based factors. In particular, we consider the Primitive Trend following Factors (PTFS) of Fung and Hsieh (2001), the Fung and Hsieh's 8-factor model including the remaining two PTFS, the SG trend indicator, Moskowitz, Ooi, and Pedersen (2012) their aggregate time series momentum factor, the Futures-based Trend-following Benchmarks Strategies of Baltas and Kosowski (2013), and our proposed adaptive time series momentum factor. Panel B analyzes the value of including the adaptive time series momentum strategy in the model proposed by Baltas and Kosowski (2013). The sample period covered in both panels equals 31/01/2000 through 30/01/2012.*, **, and *** asterisks denote significance at 90, 95, and 99% level, respectively. Robust standard errors in parentheses.

Table 5.8: Speed Factor Regressions Against Existing Risk Factors.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MKT				0.0370			
				(0.068)			
EM				-0.0328			
				(0.052)			
SMB				0.0654			
				(0.062)			
PTFSBD				-0.0581***	-0.0520***		
				(0.013)	(0.012)		
PTFSFX				-0.0135	-0.0111		
				(0.013)	(0.013)		
PTFSCOM				-0.0295**	-0.0277**		
				(0.013)	(0.013)		
PTFSIR				-0.0170***	-0.0165***		
				(0.006)	(0.006)		
PTFSSTK				-0.0557***	-0.0505***		
				(0.015)	(0.015)		
BOND				-0.0145			
				(0.009)			
CREDIT				0.5737			
				(1.010)			
BK_M			0.5603***				
			(0.048)				
BK_W			-0.5779***				
			(0.052)				
BK_D			-0.1927***				
			(0.056)				
LIQ						0.1715***	
						(0.058)	
ATSMOM	-0.0310						
	(0.092)						
MOP		0.3122***					
		(0.083)					
GVAL							0.2889*
							(0.164)
GMOM							0.6829***
							(0.168)
Constant	0.0021	-0.0020	0.0018	-0.0020	-0.0012	0.0004	-0.0012
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	215	215	215	215	215	215	215
R-squared	0.001	0.122	0.637	0.296	0.268	0.050	0.173

The tradeable speed factor returns, net of transaction costs, are regressed against existing risk factors. The adaptive TSMOM returns are net of transaction costs. The risk factors for the Fama and French (1993) and Carhart (1997) models have been downloaded from Kenneth French's website. The Pástor and Stambaugh (2003) traded liquidity factor from January 1994 to December 2014 has been obtained from Lubos Pastor's website. The Fung and Hsieh (2001) factors have been taken from David A. Hsieh's Hedge Fund Data Library. *, **, and *** asterisks denote significance at 90, 95, and 99% level, respectively. Robust standard errors in parentheses.

Table 5.9: Summary Statistics for the Speed Factor

	ATSMOM (Net of TCs)	Speed Factor (Net of TCs)	Portfolio 80/20 (net of TCs)
Min (%)	-5.91	-13.92	-6.17
Max (%)	15.22	8.89	13.17
Ann. Mean (%)	9.62	3.32	10.10
Ann. Median (%)	7.73	7.83	11.58
Ann. St. Dev. (%)	10.00	10.00	10.00
Skewness	0.71	-0.97	0.37
Kurtosis	5.61	5.74	4.21
Sharpe	0.96	0.33	1.01
Sortino	1.93	0.39	1.99
Max DD (%)	-14.13	-34.27	-13.08

Table 5.10: Asset pricing regressions on manager-based indices

	Barclay	BTOP 50	SG CTA	SG Trend	BH Syst Trend (AUM)	BH Syst Trend (ERW)
			Panel A			
ATSMOM	0.4638***	0.6541***	0.6878***	0.6842***	0.8258***	0.4398***
	(0.041)	(0.051)	(0.065)	(0.072)	(0.074)	(0.032)
Constant	0.0001	-0.0024**	-0.0018	-0.0021*	0.0002	0.0011
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.483	0.503	0.611	0.605	0.565	0.617
			Panel B			
ATSMOM	0.4518***	0.6332***	0.6881***	0.6840***	0.8200***	0.4361***
	(0.047)	(0.053)	(0.064)	(0.071)	(0.080)	(0.036)
Speed Factor	-0.1473***	-0.1219***	0.0056	-0.0031	-0.1873***	-0.1168***
	(0.037)	(0.044)	(0.052)	(0.049)	(0.062)	(0.027)
Constant	0.0007	-0.0018	-0.0019	-0.0021*	0.0008	0.0014**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.528	0.530	0.611	0.605	0.594	0.660
			Panel C			
$ATSMOM_{COM}$	0.5718***	0.5444***	0.6454***	0.6478***	0.6947***	0.5088***
	(0.105)	(0.111)	(0.142)	(0.143)	(0.147)	(0.069
$ATSMOM_{EQ}$	0.2316***	0.3742***	0.4771***	0.5908***	0.6726***	0.2872***
	(0.066)	(0.081)	(0.093)	(0.090)	(0.093)	(0.045)
$ATSMOM_{FI}$	0.6985***	1.1444***	1.2303***	1.2746***	1.5401***	0.7005***
	(0.088)	(0.125)	(0.152)	(0.154)	(0.149)	(0.070)
$ATSMOM_{FX}$	0.4982**	0.7636**	0.6464**	0.4397	0.6398*	0.3963**
	(0.214)	(0.307)	(0.297)	(0.273)	(0.349)	(0.173)
Constant	-0.0001	-0.0027**	-0.0021**	-0.0023**	-0.0000	0.0009
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.529	0.571	0.676	0.667	0.632	0.662
			Panel D			
$ATSMOM_{COM}$	0.5617***	0.5490***	0.6454***	0.6479***	0.6837***	0.5016***
	(0.109)	(0.113)	(0.142)	(0.144)	(0.151)	(0.072)
$ATSMOM_{EQ}$	0.2544***	0.3936***	0.4770***	0.5898***	0.6848***	0.2952***
	(0.068)	(0.081)	(0.091)	(0.089)	(0.095)	(0.047)
$ATSMOM_{FI}$	0.6630***	1.1086***	1.2303***	1.2743***	1.4943***	0.6705***
	(0.090)	(0.126)	(0.152)	(0.154)	(0.140)	(0.064)
$ATSMOM_{FX}$	0.4739**	0.7027**	0.6464**	0.4394	0.6500*	0.4030**
	(0.215)	(0.292)	(0.298)	(0.274)	(0.360)	(0.179)
Speed Factor	-0.1370***	-0.1154***	-0.0008	-0.0074	-0.1617***	-0.1062***
	(0.036)	(0.038)	(0.046)	(0.045)	(0.056)	(0.025)
Constant	0.0005	-0.0021**	-0.0021*	-0.0023**	0.0005	0.0012**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.564	0.591	0.676	0.667	0.653	0.697

This table shows the results of the asset pricing regressions against net returns (net of transaction costs) of the adaptive TSMOM strategy, the adaptive TSMOM strategy and speed factor, the asset class based adaptive TSMOM strategies, and the asset class based adaptive TSMOM strategy and speed factor in Panel A, B, C, and D, respectively. The asset class-based factors are adjusted to 10% annualized volatility. The dependent variables of the regressions are returns of various manager-based indices which are net of transaction costs. *, **, and *** asterisks denote significance at 90, 95, and 99% level, respectively. Robust standard errors in parentheses.

Table 5.11: Asset pricing regressions on individual Trend-Following Managed Futures

	Ann. Alpha (%)	ATSMOM COM.	ATSMOM Eq.	ATSMOM FI	ATSMOM FX	Speed Factor	Adj. R^2	Wald
Mean	0.29	0.20	0.12	0.21	0.21	-0.14	0.40	0.00
Mean* (-)	-9.55	-0.37	-0.36	-0.30	-0.18	-0.34		
Mean* (+)	4.77	0.32	0.27	0.31	0.38	0.19		
Median	0.82	0.19	0.13	0.22	0.15	-0.10	0.41	
Median* (-)	-6.56	-0.35	-0.24	-0.25	-0.12	-0.26		
Median* (+)	3.91	0.29	0.23	0.28	0.29	0.15		
t-Statistic	0.22	0.64	0.53	0.70	0.48	0.50		
t-Statistic (-)	0.06	0.03	0.04	0.02	0.03	0.42		
t-Statistic (+)	0.16	0.60	0.49	0.68	0.46	0.08		

This table shows the mean and median parameter estimates for funds that have at least one-year return history in the dataset and produce jointly significant betas at the 10% level according to the Wald test. The selection procedure leaves us with 335 funds. The rows marked with *(-) and *(+) are in reference to significant negative and positive estimates, respectively. The row named 't-Statistic' shows the share of funds that produce significant parameter estimates at 10% significance level. Similarly, the rows labeled 't-Statistic (-)' and 't-Statistic (+)' show the percentage of funds that produce, at 10% level, significant negative and positive parameters, respectively. The asset class-based adaptive TSMOM strategy factors and the speed factor are adjusted to 10% annualized volatility and net of transaction costs. The individual fund returns are net of transaction costs and gross of fees. Gross of fee returns are added assuming a 2/20 fee structure with quarterly crystallization (see Elaut, Frömmel, and Sjödin, 2015).

Table 5.12: Panel regressions on alphas

	(1)	(2)	(3)
Alpha (t-1)	0.17***	0.20***	0.21***
Log (FuM)	-0.07**	-0.12***	-0.15***
Age	0.10***	0.09***	0.09***
Fund Flow	0.00	0.00	0.00
R^2 Perfor. Regr.	3.53***	3.50***	4.26***
Com. Exp.	-1.10	-0.31	-0.17
Eq. Exp.	3.21***	3.11***	2.68***
FI Exp.	-1.47**	-1.46**	-1.31*
FX Exp.	-1.11	-1.07	-0.15
Speed Exp.	3.57***	3.69***	3.47***
Mgmt. Fee	-0.25	-0.34*	-0.29
Incent. Fee	-0.09***	-0.12***	-0.16***
ME Ratio		0.11***	0.13***
Round Turns / MUSD			0.00
No. of Obs.	2254	2007	1615
Adj. R^2	0.30	0.33	0.37

This table shows the cross-sectional analysis of the estimated alphas for 335 individual CTAs. The round turns per million USD per year and the margin-to-equity (ME ratio) statistics are not available for each CTA, therefore, in column (2) and (3) we repeat the regressions for the subset of funds for which data are available. The reported coefficients rely on a weighted least squares (WLS) panel regression that accounts for CTA period specific fixed effects. The dependent variable which is the alpha estimates from the performance regressions (see Table 5.11) is subjected to measurement errors proportional to the standard errors of the performance regressions. Therefore, in the estimation the weights are estimated standard errors of the performance regressions. The standard errors are clustered on both the specific manager and period. *, **, and *** asterisks denote significance at 90, 95, and 99% level, respectively. Robust standard errors in parentheses.

Figure 1: 3-Year Rolling Sharpe Ratio Of Rival Objective Benchmarks

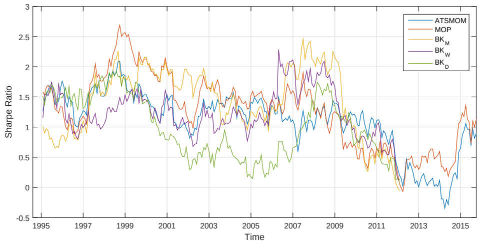


Figure 2: 3-Year Rolling Sharpe Ratios Of Manager Indices And The Benchmarks

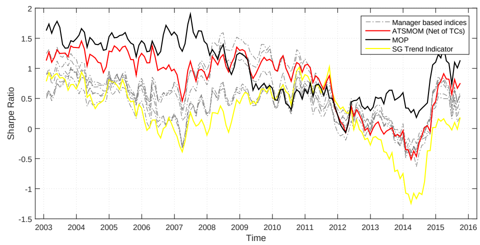


Figure 3: 60-Month Rolling Window Regression BarclayHedge (ERW) vs.CTA Benchmarks

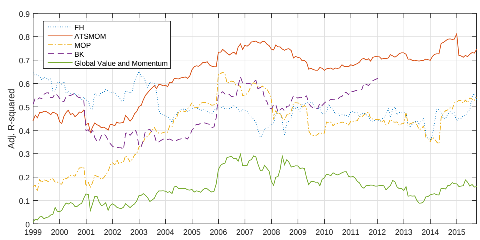


Figure 4: The Loadings Of The First Two Principal Components Of Horizon Portfolios

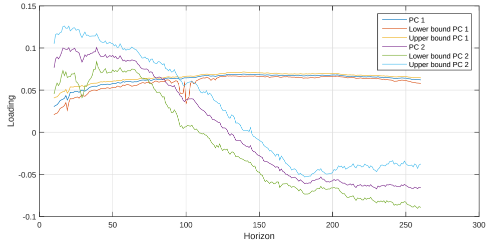


Figure 5: Portfolio performance as a function of speed factor's weight