An Examination of Liquidity and Investor Sentiment in Financial Markets

Philip Andrew Drummond

January 2019

A thesis submitted for the degree of Doctor of Philosophy of The Australian National University.

Declaration

This thesis includes working papers where I am the sole or joint author. This thesis draws on the following working papers:

1. "*Asymmetric Liquidity Persistence*".
   - Authors: Philip Andrew Drummond and Jianxin Wang.
   - Status: In preparation for journal submission.
   - Candidate contribution to research: 80%
   - Candidate contribution to authorship: 80%

   Jianxin Wang

   Jianxin Wang

2. "*Discretionary Trading Surrounding Anticipated Distraction Events: the Case of the FIFA World Cup*".
   - Authors: Philip Andrew Drummond.
   - Status: In preparation for journal submission.
   - Candidate contribution to research: 100%.
   - Candidate contribution to authorship: 100%.
3. "*Sports Sentiment and Stock Returns: An Intra-day Study*".
   - Authors: Philip Andrew Drummond.
   - Status: In preparation for journal submission.
   - Candidate contribution to research: 100%.
   - Candidate contribution to authorship: 100%.

P Drummond

Philip Andrew Drummond

Acknowledgements

I thank my supervisory panel: Raymond Liu, Jianxin Wang, Phong Ngo and Hanlin Shang. Raymond, Jianxin, Phong and Hanlin, thank you for your continuous support throughout my PhD programme. I would also like to thank Ding Ding, Jozef Drienko, Tim Higgins, Steve Sault, Alex Vadilyev and Takeshi Yamada.

I thank my fellow PhD candidates at the Research School of Finance, Actuarial Studies and Statistics. A special acknowledgment is reserved for Jan Drienko. Jan, thank you for introducing me to the SAS programming language. I also thank Souveek Halder, Jianlei Han, Bardia Khorsand, Pin-te Lin, Timothy Andrew McLennon-Smith, Duc Nguyen, Zheyao Pan, Chandler Phelps, Diego Puente Moncayo and Peter Varela.

I thank the numerous friends I have made during my PhD candidature and time spent in Canberra including: Tim Balin; Long Bui; José Armando Cobián Alvarez; Clare Cronan; Cain Edie; Rob Culling; Iván Gonzáles Gordón; Kate Herron; Sigourney Irvine; Fraser Jennison; Diego Patricio Tapia; Jeff Theys and Jack Tumminello.

I thank the friends that I have made in the College of Business and Economics football team and I thank the Food Co-op for feeding me throughout my PhD candidature.

I thank my family for their continued moral support: my mother Sonia, my father Stanley and my sister Jane. I thank my grandparents Augusta, Jean, John and Gerardo. I also thank Annabelle, Celeste, Jack, Jessica, Lisa, Lixing, Melissa, Patrick, Rafael, René and Sonia.

I reserve my greatest acknowledgement for Bec. Bec, this thesis would not have been possible without you. Thank you for your unwavering support, understanding and patience.

Discussants and Conferences

1. "Asymmetric Liquidity Persistence", with Jianxin Wang.

   I am thankful for helpful comments from Adrian Lee, seminar participants at the Australian National University, participants at the 2015 28th Australasian Finance and Banking Conference and participants at the 2015 ANU RSFAS Summer Research Camp PhD Symposium.

2. " Discretionary Trading Surround Anticipated Distraction Events: the Case of the FIFA World Cup".

   I thank Lasse Bork, James Brugler, Jan Drienko, Doug Foster, Jianfeng Shen, Wing Wah Tham and Haifeng Wu for helpful comments and suggestions. I also thank seminar participants at the Australian National University, seminar participants at the Monash University, seminar participants at the University of Queensland, participants at the 2017 7th Annual Financial Research Network (FIRN) Conference, participants at the 2017 ANU RSFAS Summer Research Camp PhD Symposium, participants at the 2017 Auckland Finance Meeting, participants at the 2017 30th Australasian Finance and Banking Conference, participants at the 2017 7th Securities Industry Research Centre of Asia-Pacific (SIRCA) Young Researchers Workshop, and participants at the 2017 30th PhD Conference in Economics and Business hosted by the University of Melbourne for helpful comments and suggestions.

3. " Sports Sentiment and Stock Returns: An Intra-day Study".

   I am thankful for helpful comments from Maren Baars and Hannes Mohrschladt, as well as participants at the 2018 Australasian Finance and Banking Conference, 2018 Monash Financial Markets Workshop and seminar participants at the Australian National University.

Scholarship Acknowledgements

I gratefully acknowledge financial support from the Australian National University, Research School of Finance, Actuarial Studies and Statistics and the Australian Postgraduate Award Programme. Throughout my candidature, I was a recipient of the following scholarships:

2013: ANU Miscellaneous Scholarship;
2014–2016: Australian Postgraduate Award; and,
2017: Postgraduate Research Scholarship.

Dedicado a mis abuelitos

Abstract

This thesis is comprised of three main chapters. The first chapter, "Asymmetric Liquidity Persistence", is based on a working paper co-authored with Jianxin Wang. In this paper, we identify a new autoregressive property of daily liquidity that contributes to sudden liquidity dry-ups. While liquidity is generally highly persistent, this persistence is conditional on past market states. Large negative returns cause liquidity persistence to initially decrease and then increase in the longer-run. We call this Asymmetric Liquidity Persistence (ALP). We show that ALP is present in both market-level and stock-level liquidity. We demonstrate that our ALP model can generate more accurate in-sample and out-of-sample liquidity estimations. According to the predictions of the Amihud (2002) liquidity premium model, our ALP model provides for a superior characterisation of the daily liquidity process.

The second chapter is titled "Discretionary Trading Surrounding Anticipated Distraction Events: the Case of the FIFA World Cup". This chapter demonstrates how anticipated market-orthogonal events can induce discretionary trading. Following Ehrmann and Jansen (2017), this study uses FIFA World Cup matches that occur during trading hours as an exogenous shock to the opportunity cost of monitoring markets. World Cup football matches have an impact on contemporaneous trading and an asynchronous impact on the rest of the trading day. In particular, when World Cup matches occur in the middle of the trading day, there is an abnormally large amount of trading between market open and kick-off time. Dollar trading volume between 120 to 90 minutes before kick-off is 23.4% of a standard deviation higher than normal levels. This is due to a temporal substitution effect whereby traders submit their orders prior to kick-off in order to avoid trading during match time. During this pre-match period, markets exhibit greater liquidity, volatility and price discovery. During matches, markets exhibit reduced liquidity, volatility and price discovery. The extraordinary market conditions that occur on match days follow the theoretical predictions of the Admati and Pfleiderer (1988) discretionary trading model.

The third chapter, "Sports Sentiment and Stock Returns: An Intra-day Study", builds upon the behavioural finance literature and in particular, the influential study of Edmans, García, and Norli (2007). Edmans et al. (2007) demonstrate that sporting results can predict overnight stock returns. The authors attribute this to a sports sentiment effect. In this thesis chapter, I demonstrate that the Edmans et al. (2007) daily sentiment effect is still present in a more recent sample of stock market data. In addition, I utilise all FIFA World Cup matches that have occurred during trading hours since 1998 to determine that there is an analogous intra-day sentiment effect. Winning full-time outcomes are associated with positive abnormal stock returns for the remainder of the trading day. Moreover, unexpected victories and victories over traditional rivals have a significant and positive marginal impact on abnormal stock returns. Using trade and quote data, this study also documents abnormal order imbalance and quote revision activity surrounding half-time match outcomes. Evidence suggests that both liquidity takers and providers are influenced by investor sentiment. Small trades exhibit the greatest sentiment effects.

Following the three main thesis chapters, I provide concluding remarks and discuss limitations and future research opportunities for each research project.

Contents

1. Asymmetric Liquidity Persistence

1.1. Abstract

We identify a new autoregressive property of daily liquidity that contributes to sudden liquidity dry-ups. While liquidity is generally highly persistent, this persistence is conditional on past market states. Large negative returns cause liquidity persistence to initially decrease and then increase in the longer-run. We call this Asymmetric Liquidity Persistence (ALP). We show that ALP is present in both market-level and stock-level liquidity. We demonstrate that our ALP model can generate more accurate in-sample and out-of-sample liquidity estimations. According to the predictions of the Amihud (2002) liquidity premium model, our ALP model provides for a superior characterisation of the daily liquidity process.

JEL Classifications: C22, C24, C58, G19.

Keywords: Liquidity, Liquidity Dry-Ups, Long-Memory, Asymmetric Persistence, Amihud Hypotheses

1.2. Introduction

The persistence of liquidity is well documented within the literature and is often a key assumption of the theoretical models of liquidity; however, persistent liquidity is also at odds with the frequently observed liquidity "dry-up" phenomenon. That is, when liquidity rapidly declines or "dries-up", often in times of crises. This dual observation suggests that while liquidity is generally persistent, its persistence is conditional on the state of the market. In this paper, we reconcile the two competing phenomena of persistent and "evaporating" liquidity by modelling the conditionality of liquidity persistence. We find that liquidity persistence is asymmetric with respect to past returns.

This study begins by confirming the baseline assumption that liquidity is persistent. We show that daily market-level and stock-level liquidity are highly persistent and display long-memory properties. Second, we demonstrate the conditionality of liquidity persistence by incorporating a threshold component into the standard heterogeneous autoregressive (HAR) model for long-memory processes. We call our model the threshold heterogeneous autoregressive model (THAR). The model involves interacting lagged liquidity with a threshold variable. The threshold variable takes the value of 1 in the case of a large positive lagged market return and the value of -1 in the case of a large negative market return. We find that the threshold-lagged liquidity interaction variables play a key role in the dynamic liquidity process. Large negative returns cause liquidity persistence to decrease in the short-term and then increase in the longer-term. We call this persistence structure asymmetric liquidity persistence (ALP).

ALP contributes significantly to the in-sample and out-of-sample explanatory power of empirical liquidity models. On average, the in-sample partial explanatory power of ALP is higher than the partial explanatory powers of both lagged returns and lagged volatility, despite returns and volatility being key determinants of contemporaneous liquidity. For stock-specific liquidity, the explanatory power of ALP is higher than that of lagged market liquidity, lagged market and stock returns, as well as lagged market volatility. To demonstrate the out-of-sample explanatory power of ALP, we perform a rolling out-of-sample forecast analysis. We find that our ALP model, the THAR model, provides superior forecasts of both market and stock liquidity.

Finally, we apply our THAR model to the asset-pricing setting. In his influential paper, Amihud (2002) theorises that if liquidity is persistent, conditional expected stock returns depend positively on expected illiquidity and "unexpected illiquidity has a negative effect on contemporaneous unexpected stock returns". We expand on the logic of Amihud's (2002) theory to conclude that current returns should be orthogonal to past and future illiquidity shocks. We estimate illiquidity residual-return cross-correlations to compare our liquidity characteristion to the Amihud (2002) theoretical prediction. We find that the THAR specification of liquidity is superior to the HAR(3) model and the AR(1) specification used by Amihud (2002) for his monthly illiquidity series. The correlation between returns and lagged illiquidity shocks converges to zero at a much faster rate when the THAR specification is used.

This paper contributes to three distinct literatures. First, this paper contributes to our understanding of the multidimensionality of liquidity dry-ups. To date, the liquidity dry-up models have focused on the causal link from returns to liquidity levels. For example, Brunnermeier and Pedersen (2009) demonstrate a funding channel by which initial negative returns lead to decreases in market liquidity; while, Kyle and Xiong (2001) and Xiong (2001) demonstrate that arbitrageurs can shift from providing liquidity to demanding liquidity after similar initial negative returns. Further, a number of studies provide empirical evidence for these dry-up theories. For example, Hameed, Kang, and Viswanathan (2010) and Nagel (2012) support the supply-side liquidity dry-up theories. Mitchell, Pedersen, and Pulvino (2007) support the margin liquidity spiral effect of Brunnermeier and Pedersen (2009). Mancini-Griffoli and Ranaldo (2011) support both the margin and loss liquidity spiral effects of Brunnermeier and Pedersen (2009). This paper not only supports these empirical results but also provides a new avenue through which liquidity dry-ups can perpetuate. We show that returns not only affect future liquidity levels but also future liquidity persistence. That is, the fundamental dynamic structure of the liquidity process changes after large positive or negative returns. Large negative returns cause liquidity persistence to initially decrease and then increase in the longer term. Our results explain, from an empirical standpoint, how liquidity can be highly persistent in normal market states and much more volatile during market downturns.

This study also contributes to the liquidity premium literature. This paper extends the empirical literature by testing the Amihud (2002) illiquidity premium hypotheses at the daily frequency and at the individual stock level. This is of key interest as many active traders have holding periods of less than one day and are concerned with individual stock liquidity. In his study, Amihud (2002) tests his theories at the annual and monthly frequencies for market liquidity, as well as portfolio level liquidity. Bekaert, Harvey, and Lundblad (2007) test the Amihud (2002) hypotheses at the monthly frequency for emerging market indices. This study also expands on the work of Amihud (2002) by developing a theoretical dynamic structure between liquidity shocks and returns. Under the assumptions of Amihud's (2002) model, returns should be orthogonal to past and future illiquidity shocks; while, by his second hypothesis, contemporaneous stock returns and unexpected illiquidity should be negatively related. This theoretical prediction provides for a new method for assessing the validity of empirical models of liquidity.

Finally, this paper contributes to the studies of the autoregressive liquidity process. In the literature, it is very common for the persistence of liquidity to be underestimated, resulting in mis-specified short-memory models of liquidity. For example: Amihud (2002) and Pástor and Stambaugh (2003) use an AR(1) model to generate illiquidity innovations; and, Acharya and Pedersen (2005) use an AR(2) model to construct illiquidity innovations. This oversight is highlighted in the aforementioned study of Bekaert et al. (2007). Bekaert et al. (2007) perform a Wald test on the liquidity residuals of their benchmark VAR(1) model. The Wald test is constructed with the null hypothesis that the first three autocorrelations are zero. They reject the null hypothesis for 11 out of the 18 emerging markets in their sample and strongly reject the null hypothesis for the joint sample. On the other hand, Wang (2013) shows that daily market illiquidity is a long-memory process. Using, Lo's 1991 modified range over standard deviation test, Wang (2013) strongly rejects the null hypothesis of no long-memory in market illiquidity for all twelve Asian equity markets in his sample. In addition, this paper confirms that daily stock level liquidity is a long-memory process, providing further justification for the use of long-memory models in liquidity.

In addition, this paper is also the first to show the conditionality of liquidity persistence. This result mirrors findings in return series (Nam, Washer, and Chu (2005); Evans and McMillan (2009)) and volatility series (McAleer and Medeiros (2008); Wang and Yang (2017)). Our analysis demonstrates the superior in-sample and out-of-sample performance of our conditional liquidity persistence model in comparison to the unconditional long-memory HAR(3) model and the short-memory AR(1) model.

The remainder of this paper is structured as follows. Section 1.3 describes the data and provides summary statistics. Section 1.4 describes our empirical model of the illiquidity process. Section 1.5 presents the empirical results of our illiquidity model. In Section 1.6 we apply our findings to the liquidity premium theory of Amihud (2002). Section 1.7 provides some concluding remarks.

1.3. Data and Summary Statistics

The market sample is composed of six major market indices: the S&P/ASX 200 Index (ASX), the Dow Jones Industrial Average Index (DOW), the FTSE 100 Index (FTSE), the NASDAQ 100 Index (NASD), the Nikkei 225 Index (NIKK) and the S&P 500 Index (SP500). The stock sample consists of the Dow 30 constituents at April 2012. The sample period is determined by the availability of daily dollar volume data. The sample period for the market indices is from the 14th of November 2001 to the 31st of December 2011. The sample period for the Dow 30 constituents is from the 14th of June 2001 to the 31st of December 2011. All data is extracted from Bloomberg.

1.3.1. Illiquidity Measure

To measure illiquidity, we use a modified version of the Amihud (2002) illiquidity ratio. The Amihud (2002) illiquidity ratio is defined as:

$$AMIHUD_t = \ln\left(\frac{|r_t|}{v_t}\right)$$

where r_t is the return of a stock or a market index on day t and v_t is the corresponding dollar volume value. The Amihud (2002) measure reflects the amount of price movement associated with a given amount of trading. The strength of the Amihud (2002) measure is that it is easy to compute and does not require intra-day data. Lou and Shu (2017) report that from 2009 to 2015, over 120 papers published in the Journal of Finance, the Journal of Financial Economics and the Review of Financial Studies use the Amihud (2002) measure in their empirical analysis. While the Amihud (2002) measure is very popular, $|r_t|$ is not an ideal measure of intra-day price movement.[1] Furthermore, Lou and Shu (2017) show that variation in the Amihud (2002) measure is mostly driven by the dollar volume component, rather than $|r_t|$.

In the volatility literature, realised volatility is promoted as a superior measure of intra-day price movement. Unfortunately, realised volatility is dependent on the availability of high-frequency returns, which are not always available for market indices. Nonetheless, most exchanges publish intra-day high and low prices for indices and stocks. High and low prices can be used to calculate the popular range measure of intra-day price movement:

$$RANGE_t = \frac{1}{2\sqrt{\ln 2}}\ln\left(\frac{P_t^H}{P_t^L}\right)$$

where P_t^H and P_t^L are daily high and low prices, respectively. The $RANGE_t$ measure has several advantages over the $|r_t|$ measure. First, the $RANGE_t$ measure directly captures intra-day price movements and is thus unlikely to be zero. Second, Parkinson (1980) shows that $RANGE_t^2$ is an unbiased estimator of the true variance of returns. Alizadeh, Brandt, and Diebold (2002) report that $RANGE_t$ is a highly efficient estimator and is robust to microstructure noise. In addition, Patton (2011) shows that the mean squared error (MSE) of $RANGE_t^2$ is approximately one-fifth of the MSE of r_t^2. Patton (2011) also shows that $RANGE_t^2$ and realised variance using six intra-day observations are just as accurate, as proxies for the true conditional variance of returns. As several leading studies of volatility dynamics favour the use of the logarithmic transformation of estimated volatility we define our intra-day price movement measure as the logarithmic transformation of

[1] For example, see Andersen and Bollerslev (1998).

$RANGE_t$:[2]

$$\sigma_t = \ln(RANGE_t)$$

The logarithmic transformation reduces the frequency of extremely large values and results in a distribution that is closer to normal.

Motivated by the volatility literature, we arrive at our modified Amihud (2002) measure of illiquidity by replacing $|r_t|$ with σ_t:

$$ILLIQ_t = \ln\left(\frac{\sigma_t}{v_t}\right). \tag{1.1}$$

For scaling purposes, v_t is measured in millions of US dollars for the stock sample, 100 billions of Yen for NIKK and billions of the domestic currency for the remaining market indices.

1.3.2. Summary Statistics

Table 1.1 presents the summary statistics for the raw illiquidity observation, $\frac{\sigma_t}{v_t}$, and the log illiquidity transformation, $ILLIQ_t$. From Table 1.1, it is apparent that there is considerable cross-sectional variation in liquidity level. This is because the indices cover different amounts of stocks, meaning that the cross-sectional dollar volume values are not directly comparable. For example, the dollar volume value of SP500 is many times larger than that of the DOW index, making the illiquidity level of the SP500 much smaller than that of the DOW index. By the same token, individual stocks have much lower dollar volume values than the indices, resulting in much higher values for the modified Amihud (2002) measure.

The raw illiquidity observations are highly positively skewed and leptokurtic, especially for ASX and FTSE. In contrast, the logarithmic transformation of the modified Amihud (2002) measure is approximately normally distributed with skewness closer to zero and kurtosis closer to 3. This is consistent with findings on the logarithmic transformation of volatility (Andersen et al. (2001a,b); Deo et al. (2006); Andersen et al. (2007); Martens et al. (2009); Wang and Yang (2009); Wang (2011)) and findings on the logarithmic transformation of liquidity depth measures (Rhee and Wang (2009); Wang (2013)).

The second last column of Table 1.1 shows the Hurst (1951) exponent for each series. In each case, the Hurst exponent is close to 1, indicating very strong long-term dependence. The last column of Table 1.1 shows Lo's (1991) modified range over standard deviation (MRS) statistic for the illiquidity series. Lo's (1991) critical values for a two-sided test of no long-memory with 99% confidence are [0.721,2.098]. Hence, for all market indices, we can strongly reject the null hypothesis of no

[2] For example, see Andersen, Bollerslev, Diebold, and Ebens (2001a); Andersen, Bollerslev, Diebold, and Labys (2001b); Deo, Hurvich, and Lu (2006); Andersen, Bollerslev, and Diebold (2007); and, Martens, Van Dijk, and De Pooter (2009).

Table 1.1
Daily Illiquidity Summary Statistics

This table presents the summary statistics for the daily raw illiquidity series, $\frac{\sigma_t}{v_t}$, and the daily log illiquidity series, $ILLIQ_t = \ln\left(\frac{\sigma_t}{v_t}\right)$, where $\sigma_t = \ln\left(\frac{1}{2\sqrt{\ln 2}}\ln\left(\frac{P_t^H}{P_t^L}\right)\right)$; P_t^H and P_t^L are the high and low prices for day t respectively; and, v_t is the corresponding dollar volume value. Hurst is the Hurst (1951) exponent. The number of lags needed to calculate the modified range over standard deviation (MRS) statistic is determined as in Lo (1991). Mean values are reported for the stock sample.

Panel A: Raw Illiquidity Series ($\frac{\sigma_t}{v_t}$)

	Mean	St Dev	Skew	Kurt	ACF(1)	Hurst	MRS
ASX	0.193	0.133	2.729	15.328	0.574	0.892	2.404
DOW	0.109	0.075	1.839	7.505	0.686	0.938	3.174
FTSE	0.184	0.153	2.485	15.195	0.708	0.921	3.530
NASD	0.095	0.054	1.700	7.769	0.641	0.924	2.948
NIKK	0.099	0.080	1.887	7.083	0.717	0.970	2.996
SP500	0.024	0.019	2.290	10.491	0.737	0.940	2.760
Stocks	4.455	3.595	2.095	10.598	0.714	0.957	3.010

Panel B: Log Illiquidity Series ($ILLIQ_t$)

	Mean	St Dev	Skew	Kurt	ACF(1)	Hurst	MRS
ASX	-1.820	0.576	0.268	3.200	0.530	0.907	2.678
DOW	-2.412	0.622	0.151	2.691	0.678	0.951	3.669
FTSE	-1.979	0.759	0.099	2.337	0.795	0.963	3.410
NASD	-2.494	0.538	-0.129	3.146	0.649	0.940	3.325
NIKK	-2.578	0.722	0.162	2.680	0.762	0.972	2.767
SP500	-3.954	0.665	0.203	2.790	0.722	0.958	3.356
Stocks	1.006	0.654	0.240	2.837	0.750	0.991	2.831

long-memory. Further, for all stocks, we reject the null hypothesis of no long-memory at the 95% level of confidence (not shown). Thus, long-run illiquidity persistence is prevalent in both the market and stock samples.

While we should not make cross-sectional comparisons of illiquidity level, we can compare skewness, kurtosis, ACF(1), Hurst and MRS across indices. On that front, it should be noted that the ASX illiquidity series appears to differ from the other series. The ASX illiquidity series has the highest skewness and kurtosis; as well as the lowest ACF(1), Hurst and MRS values.

Table 1.2 presents the correlation statistics for the daily variables of interest. As expected, the two inverse measures of liquidity depth, $AMIHUD_t$ and $ILLIQ_t$, are highly positively correlated. The correlation coefficients for $AMIHUD_t$ and $ILLIQ_t$ range from 0.490 for NASD to 0.672 for FTSE. Further, $AMIHUD_t$ and $ILLIQ_t$ share similar relationships with the other variables of interest, r_t, σ_t and v_t. For every market index, both $AMIHUD_t$ and $ILLIQ_t$ share a negative relationship with r_t, a positive relationship with σ_t and a negative relationship with v_t. The correlation coefficients corresponding to $ILLIQ_t$ are consistently of a higher magnitude than those of

Table 1.2
Daily Correlations of Illiquidity, Returns, Volatility and Turnover

This table presents the correlations of the daily variables of interest in this study. The Amihud (2002) measure, $\ln\left(\frac{|r_t|}{v_t}\right)$, is denoted by $AMIHUD_t$, where r_t is the return of a stock or a market index on day t and v_t is the corresponding dollar volume value. Volatility, σ_t, is defined as $\sigma_t = \ln\left(\frac{1}{2\sqrt{\ln 2}}\ln\left(\frac{P_t^H}{P_t^L}\right)\right)$, where P_t^H and P_t^L are the high and low prices on day t, respectively. The modified Amihud (2002) measure, $ILLIQ_t$, is defined as $\ln\left(\frac{\sigma_t}{v_t}\right)$. The stock results are from the pooled stock sample. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively.

		$AMIHUD_t$	$ILLIQ_t$	r_t	σ_t
ASX					
	$ILLIQ_t$	0.648***			
	r_t	-0.022	-0.071***		
	σ_t	0.453***	0.612***	-0.114***	
	v_t	-0.116***	-0.340***	-0.023	0.409***
DOW					
	$ILLIQ_t$	0.654***			
	r_t	-0.013	-0.038*		
	σ_t	0.519***	0.748***	-0.045**	
	v_t	-0.123***	-0.323***	-0.023	0.214***
FTSE					
	$ILLIQ_t$	0.672***			
	r_t	-0.010	-0.048**		
	σ_t	0.536***	0.716***	-0.105***	
	v_t	-0.312***	-0.576***	-0.035*	0.005
NASD					
	$ILLIQ_t$	0.490***			
	r_t	-0.008	-0.061***		
	σ_t	0.360***	0.675***	-0.074***	
	v_t	-0.107***	-0.321***	-0.037*	0.368***
NIKK					
	$ILLIQ_t$	0.547***			
	r_t	-0.035*	-0.093***		
	σ_t	0.370***	0.591***	-0.174***	
	v_t	-0.318***	-0.610***	-0.018	0.127***
SP500					
	$ILLIQ_t$	0.658***			
	r_t	-0.026	-0.054***		
	σ_t	0.533***	0.776***	-0.074***	
	v_t	-0.180***	-0.384***	-0.025	0.101***
Stocks					
	$ILLIQ_t$	0.817***			
	r_t	0.000	-0.010***		
	σ_t	0.332***	0.346***	-0.039***	
	v_t	-0.426***	-0.552***	-0.007***	0.064***

$AMIHUD_t$. This is consistent with the notion that the $ILLIQ_t$ measure is better at capturing intra-day price variation than $AMIHUD_t$.

There are several regularities in Table 1.2 that have previously been explored in the literature. As expected, returns are negatively correlated with both illiquidity measures: higher returns are associated with higher liquidity and lower returns are associated with lower liquidity. Volatility is positively and significantly correlated with both illiquidity measures. This relationship is repeatedly shown in empirical studies such as those by Weber and Rosenow (2006), Gillemot, Farmer, and Lillo (2006) and Mike and Farmer (2008). Table 1.2 also confirms the negative relation between contemporaneous realised returns and unconditional volatility, described by Christie (1982) as "part of market folklore".[3] Dollar volume is negatively and significantly correlated to both illiquidity measures. This is unsurprising as the illiquidity measures are, by construction, decreasing functions of dollar volume. The correlation coefficient between dollar volume and volatility is positive and significant at the 99% level of confidence for the stock sample and for all markets. This positive relationship is consistent with the findings of Levine and Zervos (1998), Domowitz, Glen, and Madhavan (2001) and Covitz and Downing (2007). On the other hand, the relationship between returns and dollar volume is less certain. While the correlation coefficients relating dollar volume to returns are negative, they are only statistically significant for the stock sample, NASD and FTSE. This weak negative relation is consistent with Chordia, Roll, and Subrahmanyam (2001a); Chordia, Subrahmanyam, and Anshuman (2001b) but inconsistent with the positive correlation found by Chordia, Roll, and Subrahmanyam (2002). Pástor and Stambaugh (2003) find a negative relation between contemporary returns and dollar volume for low-liquidity months and a positive relation for all other months in their sample period.

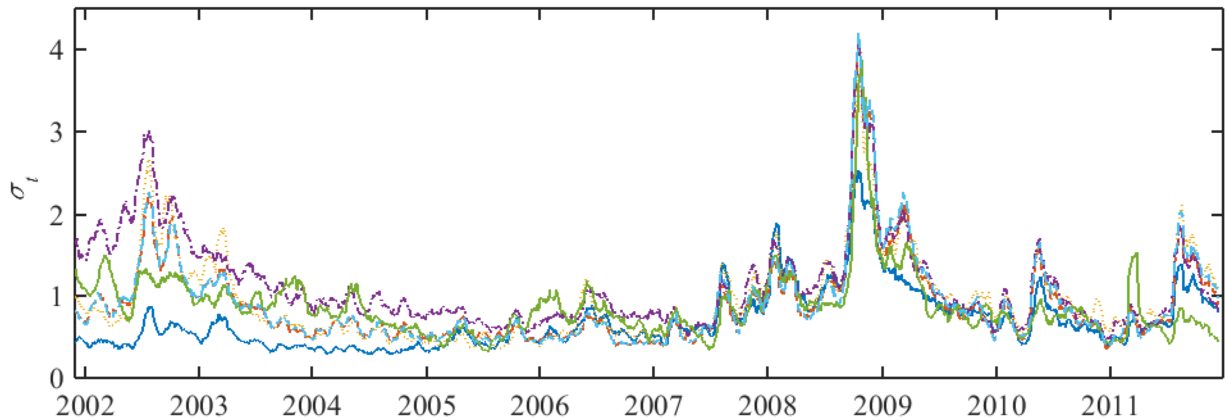
1.3.3. Seasonality Adjustments

Seasonality in liquidity is well documented. For example, Chordia, Sarkar, and Subrahmanyam (2005) demonstrate that liquidity tightness is enhanced on Mondays and Tuesdays. This study does not seek to explain seasonal variations in liquidity. Hence, it is necessary to adjust all data for seasonality.

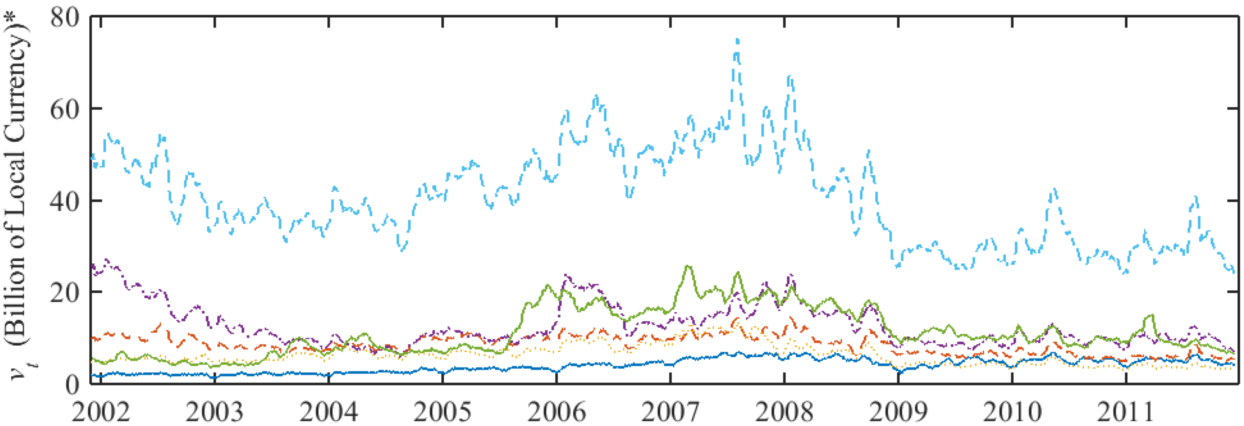
All data is adjusted using the methodology of Gallant, Rossi, and Tauchen (1992). This seasonality adjustment process maintains the mean and variance of the data, after removing all variation that is explained by the seasonal variables. The seasonal variables consist of: day-of-the-week indicator variables; month-of-the-year indicator variables; before and after non-trading day indicator variables; yearly linear and quadratic trend variables; and, a dummy variable for 2005 and onwards. The adjustments for seasonality also alleviate any concerns of non-stationarity that could be present in the non-adjusted series, seen in Figure 1.1. All results henceforth refer to the seasonally adjusted data set.

[3]The classic papers that first document this relationship include Black (1976), French, Schwert, and Stambaugh (1987) and Campbell and Hentschel (1992).

Panel A: Volatility



Panel B: Dollar Volume



*100 billion Yen for NIKK

Panel C: Illiquidity

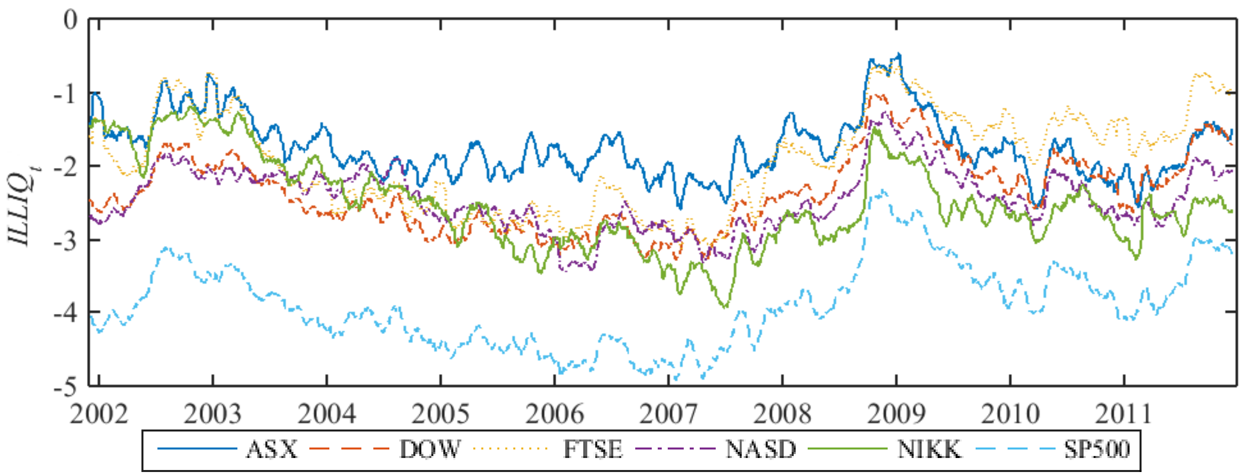


Fig. 1.1. Volatility, Dollar Volume and Illiquidity. This figure plots the daily variables of interest prior to the seasonal adjustment presented in Section 1.3.3. Volatility, σ_t, is defined as $\frac{1}{2\sqrt{\ln 2}}\ln\left(\frac{P_t^H}{P_t^L}\right)$, where P_t^H and P_t^L are the high and low prices on day t, respectively. The modified Amihud (2002) measure, $ILLIQ_t$, is defined as $\ln\left(\frac{\sigma_t}{v_t}\right)$ where v_t is dollar volume at time t.

1.4. Modelling Illiquidity Dynamics

1.4.1. Long-memory in Illiquidity

To model the illiquidity dynamics of our market indices and stocks, we require a long-memory model. There are two classes of long-memory models: (1) the fractional integration (FI) models; and, (2) the heterogeneous autoregressive (HAR) models. The FI models provide a formal framework for modelling long-memory but lack clear economic interpretation (Corsi (2009)). On the other hand, the HAR class of models are quasi long-memory models with clear economic inference. The HAR class of models were developed by Corsi (2009) and are now widely used in the volatility literature. Furthermore, Wang (2013) demonstrates that the HAR approach can be used to capture long-memory in illiquidity.

As in Wang (2013), our model resides in the HAR class of models. The HAR(3) model for illiquidity is described as:

$$ILLIQ_t^D = \alpha_0 + \sum_{k=D}^{M} \alpha_k \overline{ILLIQ}_{t-1}^k + \epsilon_t \tag{1.2}$$

where $k = D, W, M$, corresponding to "day", "week" and "month", respectively. The HAR model is adjusted to avoid overlapping lags. Hence, we define:

$$\overline{ILLIQ}_{t-1}^D \equiv ILLIQ_{t-1}^D; \qquad \overline{ILLIQ}_{t-1}^W \equiv \frac{1}{4}\sum_{i=2}^{5} ILLIQ_{t-i}^D; \qquad \overline{ILLIQ}_{t-1}^M \equiv \frac{1}{17}\sum_{i=6}^{22} ILLIQ_{t-i}^D.$$

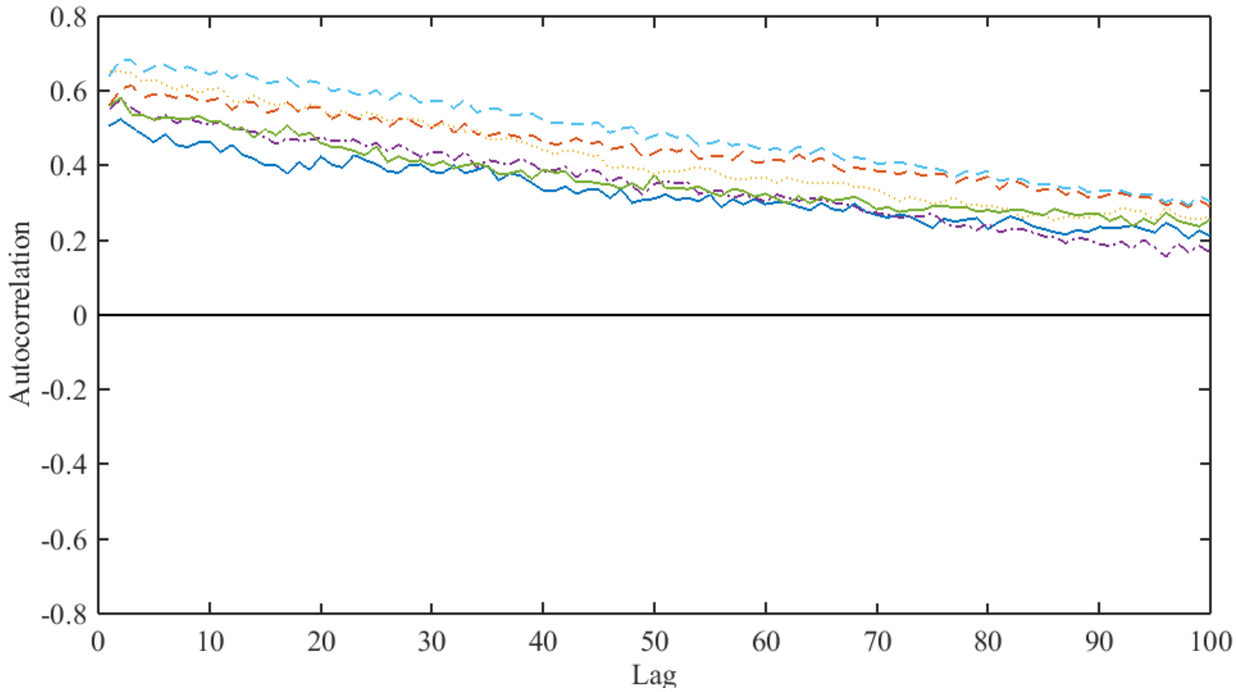
This modification of the HAR model provides an easy interpretation of the dependence of illiquidity on the illiquidity of a past period. It was first used by Patton and Sheppard (2013) and subsequently in Sévi (2014).

Figure 1.2 Panel A displays the autocorrelation function of the daily market illiquidity series. Each autocorrelation function exhibits a very slow decay, typical of long-memory processes. Figure 1.2 Panel B shows the autocorrelation function of the estimated residuals, ϵ_t, from the HAR(3) model in Equation 1.2, on the same scale as Panel A. It is apparent that the HAR(3) model captures almost all of the autocorrelation in daily illiquidity.

1.4.2. Illiquidity, Returns and Volatility

Liquidity is affected by market returns and volatility. Firstly, market returns affect investor confidence, sentiment and their ability to obtain funding to supply liquidity (Brunnermeier and Pedersen (2009)). Hameed et al. (2010) present strong evidence of a causal effect from stock returns to liquidity. Volatility, on the other hand, reflects risks from various sources, including asset fundamentals, information precision and noise trading. Higher risks increase the cost of and

Panel A: Autocorrelation functions of daily illiquidity ($ILLIQ_t$)



Panel B: Autocorrelation functions of the HAR(3)-Liq residuals (ϵ_t)

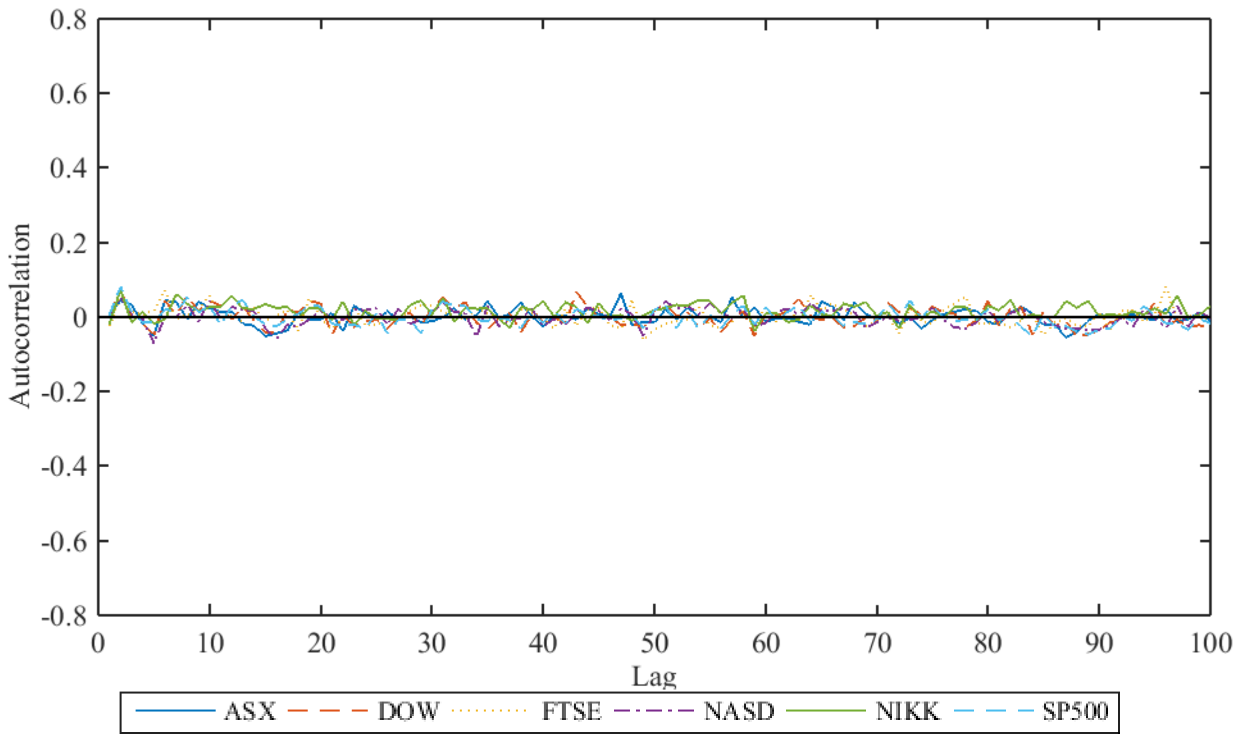


Fig. 1.2. Long-Memory in Daily Illiquidity. Panel A plots the autocorrelation functions of the modified Amihud (2002) measure, $ILLIQ_t$. The modified Amihud (2002) measure is defined as $\ln\left(\frac{\sigma_t}{v_t}\right)$ where v_t is dollar volume at time t and σ_t is the volatility at time t. Volatility, σ_t, is defined as $\frac{1}{2\sqrt{\ln 2}}\ln\left(\frac{P_t^H}{P_t^L}\right)$, where P_t^H and P_t^L are the high and low prices on day t, respectively.

the required return for supplying liquidity. Further, volatility leads to higher bid-ask spreads and lower liquidity.[4] Thus, we introduce heterogeneous return and volatility lags into the HAR framework to separately identify their liquidity impacts at different time horizons. This is similar to the heterogeneous leverage effect on volatility modelled by Corsi and Renó (2012). Thus we define the following HAR return and volatility variables in an analogous fashion to $\overline{ILLIQ}^k_{t-1}$:

$$\bar{r}^D_{t-1} \equiv r^D_{t-1}; \qquad \bar{r}^W_{t-1} \equiv \frac{1}{4}\sum_{i=2}^{5} r^D_{t-i}; \qquad \bar{r}^M_{t-1} \equiv \frac{1}{17}\sum_{i=6}^{22} r^D_{t-i}$$

$$\bar{\sigma}^D_{t-1} \equiv \sigma^D_{t-1}; \qquad \bar{\sigma}^W_{t-1} \equiv \frac{1}{4}\sum_{i=2}^{5} \sigma^D_{t-i}; \qquad \bar{\sigma}^M_{t-1} \equiv \frac{1}{17}\sum_{i=6}^{22} \sigma^D_{t-i}$$

The inclusion of these HAR variables allows us to uncover a rich dynamic structure in daily illiquidity.

1.4.3. Asymmetric Illiquidity Persistence

In studies of volatility dynamics, the dependence of volatility on past volatility, volatility persistence, is widely regarded as constant; however, McAleer and Medeiros (2008) and Wang and Yang (2017) show that the dependence of today's volatility on past volatility levels depends on the market state variables corresponding to the past volatility levels. In particular, positive returns are associated with greater volatility persistence and negative returns are associated with greater anti-persistence.

Given the strong contemporaneous relationship between liquidity and volatility and the liquidity dry-up literature that documents the effect of returns on liquidity levels, we believe that the state of the market, especially market returns, is likely to affect the persistence of illiquidity over time. Specifically, we want to explore whether and how the dependence of $ILLIQ^D_t$ on $\overline{ILLIQ}^k_{t-1}$ is affected by the corresponding HAR return values, $\bar{r}^k_{t-1}$.

Let $A_{\{k,-\}}$ and $A_{\{k,+\}}$ be positive constants and $\sigma^2_{r,k} = var\left(\bar{r}^k_{t-1}\right)$. Further, let μ_k be the mean of $\bar{r}^k_{t-1}$. We define the following threshold dummy variable to capture large returns over different time horizons:

$$\mathscr{D}^k_{t-1} = \begin{cases} -1 & \bar{r}^k_{t-1} < -A_{\{k,-\}}\sigma_{r,k} + \mu_k, \\ 0 & -A_{\{k,-\}}\sigma_{r,k} + \mu_k \leq \bar{r}^k_{t-1} \leq \mu_k + A_{\{k,+\}}\sigma_{r,k}, \\ +1 & \bar{r}^k_{t-1} > \mu_k + A_{\{k,+\}}\sigma_{r,k}. \end{cases}$$

We estimate the impact of $\mathscr{D}^k_{t-1}$ on illiquidity persistence using the following third-order threshold

[4] For example, see Wang (1999) and Wang and Yau (2000).

heterogeneous autoregressive (THAR(3)) model:[5]

$$ILLIQ_t^D = \alpha_0 + \sum_{k=D}^{M} \left[\left(\alpha_k + \beta_k \mathscr{D}_{t-1}^k \right) \overline{ILLIQ}_{t-1}^k + \theta_k \bar{r}_{t-1}^k + \lambda_k \bar{\sigma}_{t-1}^k \right] + \epsilon_t. \tag{1.3}$$

The conditional illiquidity persistence at time horizon k is measured by $\rho_t^k \equiv \alpha_k + \beta_k \mathscr{D}_{t-1}^k$, where α_k and β_k are the unconditional and conditional components, respectively. When lagged returns are large and negative ($\bar{r}_{t-1}^k < -A_{\{k,-\}}\sigma_{r,k} + \mu_k$), $\rho_t^k = \alpha_k - \beta_k$; when lagged returns are around zero ($-A_{\{k,-\}}\sigma_{r,k} + \mu_k \leq \bar{r}_{t-1}^k \leq \mu_k + A_{\{k,+\}}\sigma_{r,k}$), $\rho_t^k = \alpha_k$; and, when lagged returns are large and positive ($\bar{r}_{t-1}^k > \mu_k + A_{\{k,+\}}\sigma_{r,k}$), $\rho_t^k = \alpha_k + \beta_k$. Given illiquidity's long-memory properties, we expect $\alpha_k > 0$, for all k, as in Wang (2013). Further, given the liquidity dry-up literature and particularly, Brunnermeier and Pedersen (2009) which stipulates that liquidity dry-ups occur after recent initial losses, we expect short-run positive asymmetric persistence, $\beta_D > 0$, and a reversal in the long-run, $\beta_M < 0$. That is, after recent large negative (positive) returns illiquidity persistence decreases (increases) and we expect to find the opposite directional effect in the longer term as illiquidity reverts to more typical level of persistence. That is to say, we expect:

$$\begin{aligned}
\rho_t^D|(\mathscr{D}_{t-1}^D = -1) \leq & \rho_t^D|(\mathscr{D}_{t-1}^D = 0) \leq \rho_t^D|(\mathscr{D}_{t-1}^D = 1) \\
\rho_t^M|(\mathscr{D}_{t-1}^M = -1) \geq & \rho_t^M|(\mathscr{D}_{t-1}^M = 0) \geq \rho_t^M|(\mathscr{D}_{t-1}^M = 1).
\end{aligned}$$

For the stock sample, we consider the impacts of stock-specific and market factors on stock liquidity. This is motivated by the findings of liquidity commonality by Chordia, Roll, and Subrahmanyam (2000), Hasbrouck and Seppi (2001) and Huberman and Halka (2001). We introduce heterogeneous market illiquidity, returns and volatility lags into the model:

$$\begin{aligned}
ILLIQ_{i,t}^D = & \alpha_0 + \sum_{k=D}^{M} \left[\left(\alpha_{i,k} + \beta_{i,k} \mathscr{D}_{i,t-1}^k \right) \overline{ILLIQ}_{i,t-1}^k + \theta_{i,k} \bar{r}_{i,t-1}^k + \lambda_{i,k} \bar{\sigma}_{i,t-1}^k \right] \\
& + \sum_{k=D}^{M} \left[\alpha_{m,k} \overline{ILLIQ}_{m,t-1}^k + \theta_{m,k} \bar{r}_{m,t-1}^k + \lambda_{m,k} \bar{\sigma}_{m,t-1}^k \right] + \epsilon_t
\end{aligned} \tag{1.4}$$

where i denotes the stock and m denotes the market, represented by the S&P 500 index. Note, that for the stock sample, it is also possible to condition illiquidity persistence on *market* returns.

[5]The model in Equation 1.3 belongs to the threshold autoregressive (TAR) models of Hansen (1997). We call it the threshold heterogeneous autoregressive (THAR) model, as it captures the asymmetries in both the level and the persistence of daily illiquidity.

Thus, we can define the additional threshold variable:

$$\mathscr{D}^k_{m,t-1} = \begin{cases} -1 & \bar{r}^k_{m,t-1} < -A_{\{k,-\}}\sigma_{r,k,m} + \mu_{k,m}, \\ 0 & -A_{\{k,-\}}\sigma_{r,k,m} + \mu_{k,m} \leq \bar{r}^k_{m,t-1} \leq \mu_{k,m} + A_{\{k,+\}}\sigma_{r,k,m}, \\ +1 & \bar{r}^k_{m,t-1} > \mu_{k,m} + A_{\{k,+\}}\sigma_{r,k,m}, \end{cases}$$

where $\sigma_{r,k,m}$ is the sample standard deviation and $\mu_{k,m}$ the mean of $\bar{r}^k_{m,t-1}$, respectively. Following, Equation 1.4 can be respecified with the new threshold parameter:

$$\begin{aligned} ILLIQ^D_{i,t} =& \alpha_0 + \sum_{k=D}^{M} \left[\left(\alpha_{i,k} + \beta_{m,k}\mathscr{D}^k_{m,t-1} \right) \overline{ILLIQ}^k_{i,t-1} + \theta_{i,k}\bar{r}^k_{i,t-1} + \lambda_{i,k}\bar{\sigma}^k_{i,t-1} \right] \\ &+ \sum_{k=D}^{M} \left[\alpha_{m,k}\overline{ILLIQ}^k_{m,t-1} + \theta_{m,k}\bar{r}^k_{m,t-1} + \lambda_{m,k}\bar{\sigma}^k_{m,t-1} \right] + \epsilon_t. \end{aligned} \tag{1.5}$$

Analogous to the market sample, for the stock sample, we expect:

$$\begin{aligned} \rho^D_{i,t}|(\mathscr{D}^D_{j,t-1} = -1) \leq& \rho^D_{i,t}|(\mathscr{D}^D_{j,t-1} = 0) \leq \rho^D_{i,t}|(\mathscr{D}^D_{j,t-1} = 1) \\ \rho^M_{i,t}|(\mathscr{D}^M_{j,t-1} = -1) \geq& \rho^M_{i,t}|(\mathscr{D}^M_{j,t-1} = 0) \geq \rho^M_{i,t}|(\mathscr{D}^M_{j,t-1} = 1). \end{aligned}$$

where $j = i, m$.

1.5. Estimation and Empirical Findings

1.5.1. Model Estimation

For a given set of threshold parameters ($\{A_{\{k,-\}}$,$A_{\{k,+\}}\}$), we estimate the model using ordinary least squares, with Newey-West robust standard errors. For the stock sample, we report the average coefficients across the stocks and calculate standard errors that take into account potential cross-stock correlations. Following Hameed et al. (2010), the standard error of a cross-sectional average parameter, $\overline{\beta}$, is given by:

$$StDev\left(\overline{\beta}\right) = StDev\left(\frac{1}{30}\sum_{i=30}^{30}\beta_i\right) = \frac{1}{30}\sqrt{\sum_{i=30}^{30}\sum_{j=30}^{30}\rho_{i,j}\sigma_i\sigma_j}$$

where σ_i is the standard error of the parameter estimated for stock i and $\rho_{i,j}$ is the correlation between the regression residuals for stocks i and j. The threshold parameters $A_{\{k,+\}}$ and $A_{\{k,-\}}$ are chosen to minimise the sum of squared residuals (SSR). The range of $A_{\{k,+\}}$ and $A_{\{k,-\}}$ is $[0.1, 3]$ with a grid interval of 0.1.[6]

[6]To reduce computation time, the grid search is conducted in two stages. The optimisation is first conducted with a grid interval of 0.3, to find the first-stage optimal threshold parameters: $A^{(1)}_{\{k,+\}}$ and $A^{(1)}_{\{k,-\}}$. Then, a

1.5.2. *Market Index Results*

Table 1.3 presents the estimation results for Equation 1.3, the extended THAR(3) model. The first observation is that illiquidity exhibits an unmistakable long-run dependence: the unconditional liquidity persistence coefficients, α_k, are positive and statistically significant for all markets and all k. This is consistent with Table 1.1, which shows that both the raw illiquidity and log illiquidity observations for all markets display long-memory properties. Typically, illiquidity is most dependent on lagged monthly illiquidity, followed by lagged weekly and lagged daily illiquidity: $\hat{\alpha}_M > \hat{\alpha}_W > \hat{\alpha}_D$. Interestingly, the Asian market indices, ASX and NIKK, display a contrasting dependence structure whereby lagged weekly illiquidity is most important in determining daily illiquidity, followed by lagged monthly and lagged daily illiquidity: $\hat{\alpha}_W > \hat{\alpha}_M > \hat{\alpha}_D$. This is congruent with Wang (2013). Wang (2013) studies 12 Asian equity markets and finds that lagged weekly illiquidity is the most important autoregressive variable in determining daily illiquidity.

Table 1.3 also reveals a high degree of ALP across the sample. The coefficients of the conditional persistence variables, β_k, are statistically significant for all markets and k. The asymmetric persistence is not constant over k. In general, we observe positive ALP in the short-run and a reversal in the long-run: $\hat{\beta}_D > 0$ and $\hat{\beta}_M < 0$; with the exception of DOW. For weekly conditional persistence, β_W, we observe positive asymmetric persistence ($\hat{\beta}_W > 0$), except for the ASX which reverses to negative ALP at a faster rate than the other indices. The ASX's more rapid reversal in illiquidity persistence is also much more pronounced than the other markets: $\hat{\beta}_{ASX,W} = -0.115$ is of a substantially larger magnitude than all other $\hat{\beta}_k$ coefficients. Hence, the ASX differs in terms of distribution (Table 1.1) and in terms of heterogeneous conditional illiquidity persistence. To that end, the ASX also displays the lowest adjusted R-squared in Table 1.3.

Panel B of Table 1.3 reveals a strictly negative effect of lagged returns on daily illiquidity. The estimated lagged return coefficients, $\hat{\theta}_k$, are uniformly negative and significant at the 5% level, except $\hat{\theta}_M = -0.056$, for DOW, which is statistically insignificant. Thus, lagged positive returns increase current liquidity. This is consistent with theoretical models and the empirical results of Hameed et al. (2010), which show that negative market returns increase market illiquidity. The dynamic effect between liquidity and lagged returns is fairly consistent across the sample. For all indices except DOW, $|\hat{\theta}_M| > |\hat{\theta}_W| > |\hat{\theta}_D|$.

In contrast to the positive contemporaneous relationship between illiquidity and volatility, Table 1.3 shows a negative relationship between lagged daily price volatility and daily illiquidity: all estimates of the lagged daily volatility coefficient, λ_D, are negative, with four out of the six being significantly different from zero at the 90% level of confidence. Table 1.3 Panel B also reveals that lagged long-run volatility increases daily illiquidity: all significant estimates of λ_W and λ_M are greater than zero. This mirrors Fung and Patterson (1999), whom find a dynamic reversal

second traversing grid search is conducted with an interval of 0.1 over ranges $\left[A^{(1)}_{\{k,+\}} - 0.3, A^{(1)}_{\{k,+\}} + 0.3\right]$ and $\left[A^{(1)}_{\{k,-\}} - 0.3, A^{(1)}_{\{k,-\}} + 0.3\right]$, to find $A_{\{k,+\}}$ and $A_{\{k,-\}}$, respectively.

Table 1.3
Daily Market Illiquidity Dynamics

This table reports the estimation results for the following threshold regression:

$$ILLIQ_t^D = \alpha_0 + \sum_{k=D}^{M} \left[\left(\alpha_k + \beta_k \mathscr{D}_{t-1}^k\right) \overline{ILLIQ}_{t-1}^k + \theta_k \bar{r}_{t-1}^k + \lambda_k \bar{\sigma}_{t-1}^k \right] + \epsilon_t \tag{1.3}$$

where $ILLIQ_t$ is the modified Amihud (2002) measure defined as $\ln\left(\frac{\sigma_t}{v_t}\right)$ where v_t is dollar volume at time t and σ_t is defined as $\sigma_t = \ln\left(\frac{1}{2\sqrt{\ln 2}} \ln\left(\frac{P_t^H}{P_t^L}\right)\right)$, where P_t^H and P_t^L are the high and low prices on day t, respectively. The market index return at time t is denoted by r_t. The independent variables are heterogeneous autoregressive lags corresponding to the $ILLIQ_t$, r_t and σ_t variables and superscript $k = D, W, M$, corresponding to "day", "week" and "month", respectively. For $k = D, W$ and M, the threshold parameter is defined as:

$$\mathscr{D}_{t-1}^k = \begin{cases} -1 & \bar{r}_{t-1}^k < -A_{\{k,-\}}\sigma_{r,k} + \mu_k, \\ 0 & -A_{\{k,-\}}\sigma_{r,k} + \mu_k \leq \bar{r}_{t-1}^k \leq \mu_k + A_{\{k,+\}}\sigma_{r,k}, \\ +1 & \bar{r}_{t-1}^k > \mu_k + A_{\{k,+\}}\sigma_{r,k}, \end{cases}$$

where $\sigma_{r,k}$ is the sample standard deviation and μ_k is the mean of $\bar{r}_{t-1}^k$, respectively. The threshold parameters $A_{\{k,+\}}$ and $A_{\{k,-\}}$ are chosen to minimise the sum of squared residuals (SSR). The range of $A_{\{k,+\}}$ and $A_{\{k,-\}}$ is $[0.1, 3]$ with a grid interval of 0.1. The t-statistics are reported in italics. Newey-West robust standard errors are used throughout. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively.

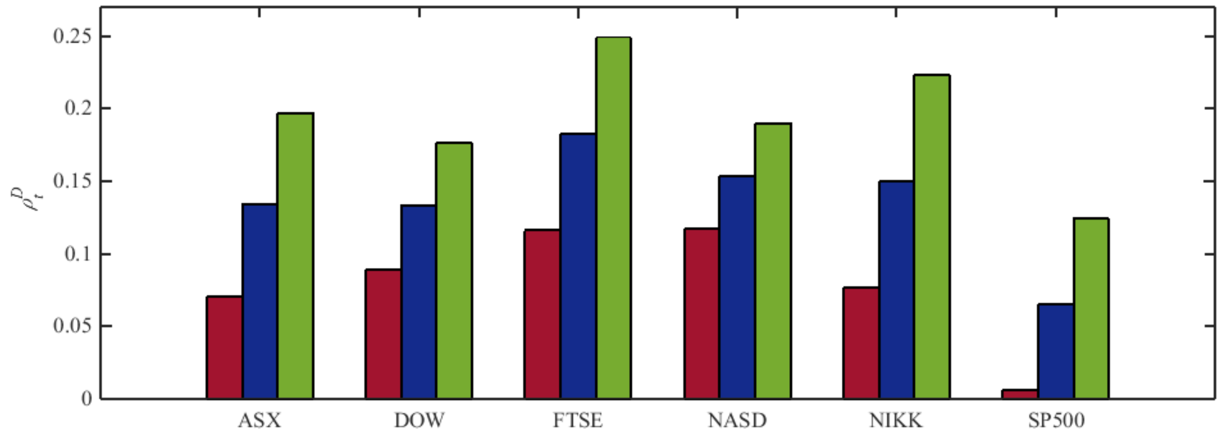
Panel A: Illiquidity Dynamics

	α_D	α_W	α_M	β_D	β_W	β_M	$A_{\{D,-\}}$	$A_{\{D,+\}}$	$A_{\{W,-\}}$	$A_{\{W,+\}}$	$A_{\{M,-\}}$	$A_{\{M,+\}}$
ASX	0.134***	0.336***	0.236***	0.063***	-0.115***	-0.061***	2.3	0.2	2.9	1.5	1.5	1.1
	3.00	*4.95*	*3.33*	*3.23*	*-2.74*	*-3.01*						
DOW	0.133***	0.199**	0.552***	0.044***	0.064***	0.031*	0.1	0.6	2.7	0.6	0.3	2.9
	2.58	*2.16*	*4.48*	*4.02*	*3.39*	*1.79*						
FTSE	0.183***	0.297***	0.413***	0.066***	0.055***	-0.035**	0.2	0.7	0.4	1.0	2.5	0.6
	4.87	*3.48*	*4.54*	*3.59*	*2.83*	*-1.63*						
NASD	0.154***	0.338***	0.425***	0.036***	0.026**	-0.033***	3.0	0.2	0.6	0.3	2.9	0.9
	3.46	*4.51*	*6.10*	*3.50*	*2.53*	*-2.84*						
NIKK	0.150***	0.396***	0.181***	0.073***	0.036*	-0.060***	2.0	0.3	1.0	0.4	3.0	0.6
	2.99	*4.12*	*2.91*	*3.18*	*1.76*	*-3.02*						
SP500	0.065**	0.320***	0.573***	0.059***	0.030***	-0.056***	2.6	0.5	0.6	0.6	2.6	1.6
	2.17	*2.87*	*4.69*	*4.40*	*3.32*	*-2.66*						

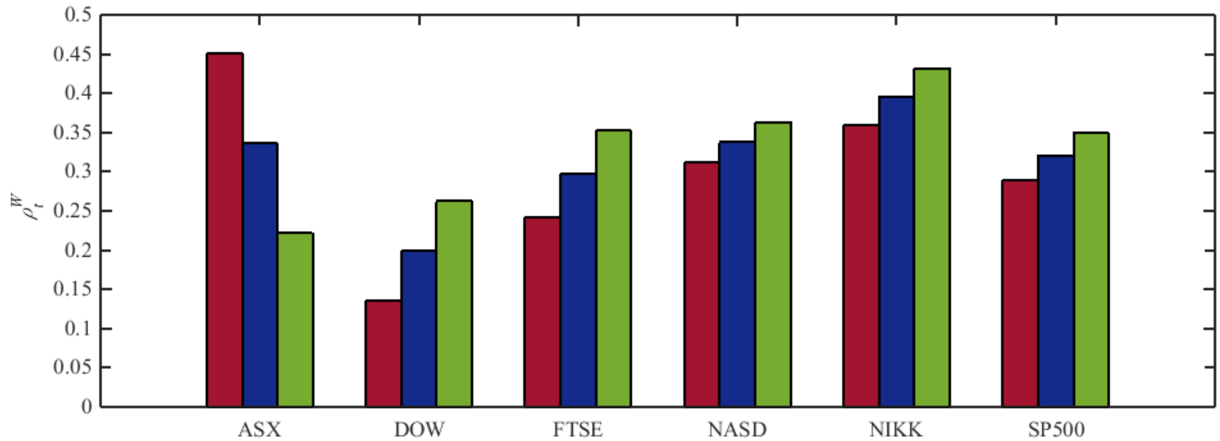
Panel B: Impacts of Return and Volatility on Illiquidity

	θ_D	θ_W	θ_M	λ_D	λ_W	λ_M	α_0	$\bar{R}^2$
ASX	-0.042***	-0.155***	-0.247***	-0.058**	0.018	0.167**	-0.513***	0.316
	-3.27	*-5.75*	*-3.60*	*-1.93*	*0.34*	*2.39*	*-4.05*	
DOW	-0.031***	-0.083***	-0.056	-0.131***	0.247***	-0.022	-0.354***	0.395
	-3.19	*-3.61*	*-0.90*	*-3.24*	*3.37*	*-0.27*	*-2.88*	
FTSE	-0.027***	-0.085***	-0.249***	-0.116***	0.096	0.135	-0.344***	0.390
	-2.59	*-2.90*	*-3.68*	*-2.41*	*0.93*	*1.36*	*-3.57*	
NASD	-0.026***	-0.054***	-0.129***	-0.129***	0.006	0.106*	-0.229***	0.460
	-4.68	*-3.29*	*-4.79*	*-3.36*	*0.09*	*1.89*	*-2.75*	
NIKK	-0.033***	-0.082**	-0.247***	-0.092	-0.120	0.452***	-0.708***	0.362
	-2.78	*-2.20*	*-3.57*	*-1.34*	*-1.20*	*4.46*	*-5.48*	
SP500	-0.026**	-0.054**	-0.215***	-0.055	0.083	-0.005	-0.204	0.443
	-2.39	*-2.33*	*-3.19*	*-1.35*	*0.95*	*-0.05*	*-0.85*	

Panel A: Daily Conditional Market Illiquidity Persistence (ρ_t^D)



Panel B: Weekly Conditional Market Illiquidity Persistence (ρ_t^W)



Panel C: Monthly Conditional Market Illiquidity Persistence (ρ_t^M)

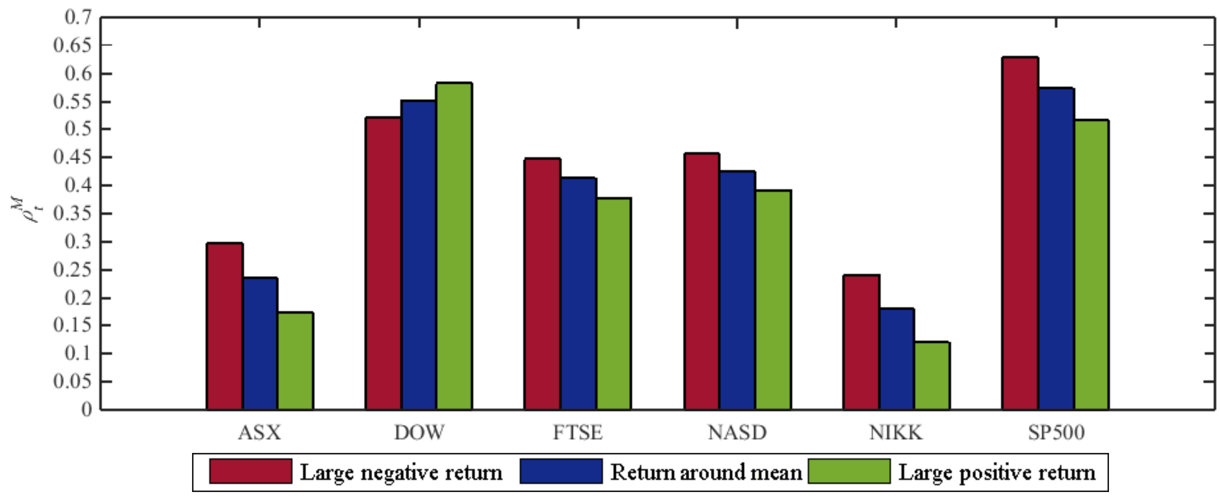


Fig. 1.3. Conditional Market Illiquidity Persistence. This figure plots the conditional illiquidity persistence ($\rho_t^k \equiv \alpha_k + \beta_k \mathscr{D}_{t-1}^k$) values for each market index and value of k, where $k = D, W, M$. The green bars correspond to the ρ_t^k values where $\bar{r}_{t-1}^k > \mu_k + A_{\{k,+\}}\sigma_{r,k}$ and $\mathscr{D}_{t-1}^k = 1$; the blue bars correspond to the ρ_t^k values where $-A_{\{k,-\}}\sigma_{r,k} + \mu_k \leq \bar{r}_{t-1}^k \leq \mu_k + A_{\{k,+\}}\sigma_{r,k}$ and $\mathscr{D}_{t-1}^k = 0$; and, the red bars correspond to the ρ_t^k values where $\bar{r}_{t-1}^k < -A_{\{k,-\}}\sigma_{r,k} + \mu_k$ and $\mathscr{D}_{t-1}^k = -1$. The α_k and β_k values are taken from the estimation of Equation 1.3, presented in Table 1.3.

Table 1.4
Conditional Market Illiquidity Persistence and Frequency of Extreme Days

This table displays the conditional illiquidity persistence ($\rho_t^k \equiv \alpha_k + \beta_k \mathscr{D}_{t-1}^k$) values for each market index and value of k, where $k = D, W, M$. The α_k and β_k values are taken from the estimation of Equation 1.3, presented in Table 1.3. The table also gives the frequency of each $\mathscr{D}_{t-1}^k$ value, denoted in italics and given as a percentage. The far right column shows the percentage change in ρ_t^k from $\mathscr{D}_{t-1}^k = -1$ to $\mathscr{D}_{t-1}^k = 1$.

Panel A: Daily Conditional Illiquidity Persistence (ρ_t^D)

		(a) Large Negative Return ($\mathscr{D}_{t-1}^D = -1$)	Return Around Mean ($\mathscr{D}_{t-1}^D = 0$)	(b) Large Positive Return ($\mathscr{D}_{t-1}^D = 1$)	% Δ from (a) to (b)
ASX	$\hat{\rho}_t^D$	0.070	0.134	0.197	179.9%
	Frequency	*1.9%*	*57.2%*	*40.9%*	
DOW	$\hat{\rho}_t^D$	0.089	0.133	0.177	97.9%
	Frequency	*40.8%*	*40.1%*	*19.1%*	
FTSE	$\hat{\rho}_t^D$	0.116	0.183	0.249	114.2%
	Frequency	*37.4%*	*45.6%*	*17.0%*	
NASD	$\hat{\rho}_t^D$	0.117	0.154	0.190	61.8%
	Frequency	*0.8%*	*59.1%*	*40.1%*	
NIKK	$\hat{\rho}_t^D$	0.077	0.150	0.223	190.6%
	Frequency	*2.3%*	*63.1%*	*34.6%*	
SP500	$\hat{\rho}_t^D$	0.006	0.065	0.124	1954.9%
	Frequency	*1.2%*	*75.9%*	*22.8%*	

Panel B: Weekly Conditional Illiquidity Persistence (ρ_t^W)

		(a) Large Negative Return ($\mathscr{D}_{t-1}^W = -1$)	Return Around Mean ($\mathscr{D}_{t-1}^W = 0$)	(b) Large Positive Return ($\mathscr{D}_{t-1}^W = 1$)	% Δ from (a) to (b)
ASX	$\hat{\rho}_t^W$	0.451	0.336	0.222	-50.9%
	Frequency	*1.1%*	*94.7%*	*4.2%*	
DOW	$\hat{\rho}_t^W$	0.135	0.199	0.263	93.8%
	Frequency	*1.2%*	*75.9%*	*22.9%*	
FTSE	$\hat{\rho}_t^W$	0.242	0.297	0.353	45.8%
	Frequency	*28.2%*	*61.0%*	*10.8%*	
NASD	$\hat{\rho}_t^W$	0.312	0.338	0.363	16.5%
	Frequency	*23.5%*	*38.5%*	*38.0%*	
NIKK	$\hat{\rho}_t^W$	0.360	0.396	0.432	20.1%
	Frequency	*12.5%*	*54.9%*	*32.6%*	
SP500	$\hat{\rho}_t^W$	0.290	0.320	0.349	20.5%
	Frequency	*22.5%*	*54.2%*	*23.3%*	

Panel C: Monthly Conditional Illiquidity Persistence (ρ_t^M)

		(a) Large Negative Return ($\mathscr{D}_{t-1}^M = -1$)	Return Around Mean ($\mathscr{D}_{t-1}^M = 0$)	(b) Large Positive Return ($\mathscr{D}_{t-1}^M = 1$)	% Δ from (a) to (b)
ASX	$\hat{\rho}_t^M$	0.297	0.236	0.174	-41.4%
	Frequency	*7.8%*	*82.7%*	*9.5%*	
DOW	$\hat{\rho}_t^M$	0.522	0.552	0.583	11.8%
	Frequency	*30.5%*	*69.3%*	*0.2%*	
FTSE	$\hat{\rho}_t^M$	0.448	0.413	0.378	-15.6%
	Frequency	*1.8%*	*73.4%*	*24.8%*	
NASD	$\hat{\rho}_t^M$	0.458	0.425	0.392	-14.4%
	Frequency	*1.3%*	*83.7%*	*15.0%*	
NIKK	$\hat{\rho}_t^M$	0.241	0.181	0.120	-50.0%
	Frequency	*0.7%*	*74.1%*	*25.2%*	
SP500	$\hat{\rho}_t^M$	0.629	0.573	0.518	-17.7%
	Frequency	*1.6%*	*96.4%*	*2.0%*	

relationship between lagged volatility and market depth.

Figure 1.3 provides a graphical representation of the market illiquidity persistence results in Table 1.3 Panel A. Panel A of Figure 1.3 shows that for all market indices, we observe positive and significant daily ALP:

$$\hat{\rho}_t^D|(\mathscr{D}_{t-1}^D = -1) \leq \hat{\rho}_t^D|(\mathscr{D}_{t-1}^D = 0) \leq \hat{\rho}_t^D|(\mathscr{D}_{t-1}^D = 1).$$

SP500 displays the lowest daily illiquidity persistence and the largest degree of asymmetric persistence: daily illiquidity persistence increases by 1954.9%, when comparing observations following a large negative return day (1.2% of observations) to a large positive return day (22.8% of observations). When comparing a "normal" return day (75.9% of observations), to a large negative return day (1.2% of observations), daily conditional illiquidity persistence decreases by 90.7%. This describes how the highly persistent dynamic structure of illiquidity can breakdown in times of market downturn. Panel B shows that the ASX is the only market index to exhibit negative ALP at the weekly lag. ASX also has the second highest degree of weekly asymmetric illiquidity persistence, with a -50.9% change in persistence following a large negative return week (1.1% of observations) to a large positive return week (4.2% of observations). The equivalent figure for DOW is 93.8%. Figure 1.3 Panel C shows that for all indices except DOW, we observe negative ALP in the long-run:

$$\hat{\rho}_t^M|(\mathscr{D}_{t-1}^M = -1) \geq \hat{\rho}_t^M|(\mathscr{D}_{t-1}^M = 0) \geq \hat{\rho}_t^M|(\mathscr{D}_{t-1}^M = 1).$$

The DOW index displays modest positive ALP at the monthly lag. Table 1.4 complements Figure 1.3 in that it details the frequency with which large negative or large positive returns are observed.

1.5.3. Stock Results

Table 1.5 presents the mean estimated regression coefficients for equations 1.4 and 1.5. Panels A and B of Table 1.5 display the stock-specific dependencies of stock illiquidity. The results presented in Table 1.5 panels A and B mirror the market illiquidity dynamics presented in Table 1.3. Firstly, like market illiquidity, stock illiquidity exhibits very strong unconditional illiquidity persistence. All $\alpha_{k,i}$ variables are positive and highly significant. Further, all $\theta_{i,k}$ coefficients are negative and statistically different from zero, indicating, that negative returns lead to decreased levels of liquidity, as in Hameed et al. (2010). Analogous to Table 1.3, monthly returns have the greatest influence over liquidity levels: $|\bar{\bar{\theta}}_{i,M}| > ||\bar{\bar{\theta}}_{i,W}| > ||\bar{\bar{\theta}}_{i,D}|$. Similar to Table 1.3, the estimated $\lambda_{i,k}$ coefficients indicate a unique dynamic relationship between volatility and illiquidity. The results indicate that stock volatility has an initial positive effect on liquidity and a negative impact on liquidity in the long-run. That is, $\lambda_{i,D} < 0$ and $\lambda_{i,M} > 0$.

Table 1.5 Panel A also indicates that stock illiquidity, like market illiquidity, displays ALP; however, in this instance, illiquidity persistence is sensitive to past stock *and* market returns. For

Table 1.5
Daily Stock Illiquidity Dynamics

This table reports the stock regression results for the equation:

$$
\begin{aligned}
ILLIQ_{i,t}^{D} =& \alpha_0 + \sum_{k=D}^{M} \left[\left(\alpha_{i,k} + \beta_{j,k} \mathscr{D}_{j,t-1}^{k} \right) \overline{ILLIQ}_{i,t-1}^{k} + \theta_{i,k} \bar{r}_{i,t-1}^{k} + \lambda_{i,k} \bar{\sigma}_{i,t-1}^{k} \right] \\
&+ \sum_{k=D}^{M} \left[\alpha_{m,k} \overline{ILLIQ}_{m,t-1}^{k} + \theta_{m,k} \bar{r}_{m,t-1}^{k} + \lambda_{m,k} \bar{\sigma}_{m,t-1}^{k} \right] + \epsilon_t
\end{aligned}
$$

where $ILLIQ_{i,t}$ is the modified Amihud (2002) measure defined as $\ln\left(\frac{\sigma_{i,t}}{v_{i,t}}\right)$ where $v_{i,t}$ is dollar volume at time t for stock i and $\sigma_{i,t}$ is defined as $\sigma_{i,t} = \ln\left(\frac{1}{2\sqrt{\ln 2}} \ln\left(\frac{P_{i,t}^{H}}{P_{i,t}^{L}}\right)\right)$, where $P_{i,t}^{H}$ and $P_{i,t}^{L}$ are the high and low prices on day t for stock i, respectively. The subscript m denotes the market. The S&P 500 index is used to proxy the market. Stock returns at time t are denoted by $r_{i,t}$ and market returns are denoted by $r_{m,t}$, respectively. The independent variables are heterogeneous autoregressive lags corresponding to the $ILLIQ_{i,t}$, $r_{i,t}$, $\sigma_{i,t}$, $ILLIQ_{m,t}$, $r_{m,t}$ and $\sigma_{m,t}$ variables and superscript $k = D, W, M$, corresponding to "day", "week" and "month", respectively. For $k = D, W$ and M, the threshold parameters are defined as:

$$
\mathscr{D}_{j,t-1}^{k} = \begin{cases} -1 & \bar{r}_{j,t-1}^{k} < -A_{\{k,-\}} \sigma_{r,k,j} + \mu_{k,j}, \\ 0 & -A_{\{k,-\}} \sigma_{r,k,j} + \mu_{k,j} \leq \bar{r}_{j,t-1}^{k} \leq \mu_{k,j} + A_{\{k,+\}} \sigma_{r,k,j}, \\ +1 & \bar{r}_{j,t-1}^{k} > \mu_{k,j} + A_{\{k,+\}} \sigma_{r,k,j}, \end{cases}
$$

where $\sigma_{r,k,j}$ is the sample standard deviation of and $\mu_{k,j}$ is the mean of $\bar{r}_{j,t-1}^{k}$, where $j = i, m$, respectively. The average coefficients are reported with their associated Hameed et al. (2010) *t*-statistics in italics. The mean $A_{\{k,-\}}$, $A_{\{k,+\}}$ and R^2 values are reported. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively.

Panel A: Illiquidity Dynamics

	Equation 1.4	Equation 1.5
α_0	0.829***	0.823***
	6.07	*6.08*
$\alpha_{i,D}$	0.191***	0.193***
	14.59	*14.47*
$\alpha_{i,W}$	0.254***	0.255***
	13.14	*13.28*
$\alpha_{i,M}$	0.222***	0.222***
	9.79	*9.83*
$\beta_{i,D}$	0.057***	
	3.49	
$\beta_{i,W}$	0.047**	
	2.88	
$\beta_{i,M}$	0.004	
	0.22	
$\beta_{m,D}$		0.038**
		2.59
$\beta_{m,W}$		0.052***
		3.74
$\beta_{m,M}$		-0.048**
		-3.03

Table 1.5
(continued)

Panel B: Impact of Return and Volatility on Illiquidity		
	Equation 1.4	Equation 1.5
$\theta_{i,D}$	-0.020***	-0.014**
	-4.03	*-3.32*
$\theta_{i,W}$	-0.041***	-0.028***
	-3.89	*-3.36*
$\theta_{i,M}$	-0.049*	-0.044*
	-1.97	*-2.19*
$\lambda_{i,D}$	-0.138***	-0.141***
	-8.62	*-8.82*
$\lambda_{i,W}$	-0.025	-0.024
	-0.91	*-0.88*
$\lambda_{i,M}$	0.139***	0.141***
	3.47	*3.51*

Panel C: Market Variables		
	Equation 1.4	Equation 1.5
$\alpha_{m,D}$	-0.009	-0.009
	-0.67	*-0.69*
$\alpha_{m,W}$	-0.126***	-0.128***
	-3.44	*-3.48*
$\alpha_{m,M}$	0.266***	0.266***
	5.92	*5.92*
$\theta_{m,D}$	-0.021***	-0.025***
	-4.29	*-4.06*
$\theta_{m,W}$	-0.074***	-0.086***
	-6.30	*-5.98*
$\theta_{m,M}$	-0.049	-0.026
	-1.63	*-0.73*
$\lambda_{m,D}$	0.023	0.024
	1.47	*1.48*
$\lambda_{m,W}$	0.169***	0.167***
	4.49	*4.45*
$\lambda_{m,M}$	-0.093	-0.093
	-1.85	*-1.84*

Panel D: Other		
	Equation 1.4	Equation 1.5
$A_{\{D,-\}}$	1.417	1.427
$A_{\{D,+\}}$	1.617	1.327
$A_{\{W,-\}}$	1.797	1.790
$A_{\{W,+\}}$	1.313	1.410
$A_{\{M,-\}}$	1.353	1.573
$A_{\{M,+\}}$	1.613	1.737
$\bar{R}^2$	0.331	0.331

Equation 1.4 we observe positive short-run ALP with $\bar{\bar{\beta}}_{i,D}$ and $\bar{\bar{\beta}}_{i,W}$ being positive and significantly different from zero. Thus, large lagged daily and weekly returns have the effect of decreasing daily illiquidity persistence. Similarly, for Equation 1.5, we find that $\bar{\bar{\beta}}_{m,D}$ and $\bar{\bar{\beta}}_{m,W}$ are positive and significant. Further, like the market sample, we observe a reversal in ALP in the longer run: $\bar{\bar{\beta}}_{m,M} < 0$. This means that *stock* liquidity persistence decreases after large negative *market* returns. This is congruent with the liquidity "dry-up" theories, such as that of Brunnermeier and Pedersen (2009). Brunnermeier and Pedersen (2009) illustrate how negative market returns can cause stock liquidity to drastically decrease as speculators are forced to reduce their positions in reaction to worsening funding constraints. Moreover, examining the mean estimated $\theta_{m,k}$ coefficients, we observe that market returns have a stronger negative impact on stock liquidity than stock-specific returns: at the daily frequency $|\theta_{m,D}| > |\theta_{i,D}|$; and, at the weekly frequency $|\theta_{m,W}| > |\theta_{i,W}|$. This supports the notion that liquidity dry-ups are caused by market-wide declines and propagated by traders invested in multiple assets.

Other results in Table 1.5 Panel C show that stock illiquidity is negatively correlated to lagged weekly market illiquidity ($\bar{\bar{\alpha}}_{m,W} < 0$) and positively related to lagged monthly market illiquidity ($\bar{\bar{\alpha}}_{m,M} > 0$). This provides evidence of a lead-lag relationship between stock liquidity and market liquidity. As, the Dow 30 stocks are large stocks, this result is consistent with the large to small lead-lag liquidity relationship found by Chordia, Sarkar, and Subrahmanyam (2011). Finally, both mean estimated $\lambda_{m,W}$ values are positive and statistically significant, indicating that daily stock illiquidity is positively related to lagged weekly market volatility.

1.5.4. Explanatory Power

This section explores the relative explanatory powers of unconditional liquidity dependence, ALP, returns and volatility to the daily illiquidity process. This is achieved by performing a partial R-squared analysis whereby each category of lagged variables is grouped across k. For example, $PR_{i,ILLIQ}$ represents that partial R-squared statistics for $\overline{ILLIQ}^k_{i,t-1}$, for $k = D, W, M$; while, $PR_{i,ALP}$ represents that partial R-squared statistics for the $\mathscr{D}^k_{i,t-1} \cdot \overline{ILLIQ}^k_{i,t-1}$ interaction variables. Table 1.6, shows the partial R-squared statistics for equations 1.3, 1.4 and 1.5. Unconditional persistence is by far the greatest contributor to detrended daily illiquidity. At the market level, the ASX has the lowest unconditional illiquidity dependence with $PR_{ASX,ILLIQ} = 6.03\%$; while NASD has the highest unconditional illiquidity dependence with $PR_{NASD,ILLIQ} = 28.15\%$. The average market PR_{ILLIQ} value is 13.01%. Unconditional persistence is also the largest contributor to stock illiquidity. The mean partial R-squared statistic for the $\overline{ILLIQ}^k_{i,t-1}$ variables and Equation 1.5 is 11.97%. This highlights the importance of long-memory in illiquidity.

Perhaps the most striking result from Table 1.6 is that the explanatory power of ALP is of a similar order of magnitude to the explanatory powers of the well-established contemporaneous determinants of illiquidity: returns and volatility. At the market level, the mean PR_{ALP} value is 1.41%; while, the mean PR_r and PR_σ values are 0.98% and 0.79%, respectively. In Panel B,

Table 1.6
Partial R-Squared Statistics

This table reports the partial R-squared statistics for the THAR analyses of illiquidity. For the market indices, the partial R-squared statistics relating to Equation 1.3 are reported. For example, $PR_{i,ILLIQ}$ represents that partial R-squared statistics for $\overline{ILLIQ}^{k}_{i,t-1}$, for $k = D, W, M$; while, $PR_{i,ALP}$ represents that partial R-squared statistics for the $\mathscr{D}^{k}_{i,t-1} \cdot \overline{ILLIQ}^{k}_{i,t-1}$ interaction variables. For the stock sample, the average partial R-squared statistics for equations 1.4 and 1.5 are reported.

Panel A: Market Index Sample (Equation 1.3)

	PR_{ILLIQ}	PR_{ALP}	PR_r	PR_σ
ASX	6.03%	1.08%	1.78%	0.43%
DOW	7.29%	1.56%	0.53%	1.12%
FTSE	13.11%	1.25%	0.59%	0.55%
NASD	28.15%	1.45%	1.43%	0.96%
NIKK	12.90%	1.40%	0.91%	1.50%
SP500	10.58%	1.70%	0.64%	0.19%

Panel B: Stock Sample (Equations 1.4 and 1.5)

	Equation	$PR_{i,ILLIQ}$	$PR_{i,ALP}$	$PR_{m,ALP}$	$PR_{i,r}$	$PR_{i,\sigma}$	$PR_{m,ILLIQ}$	$PR_{m,r}$	$PR_{m,\sigma}$
AA	4	19.69%	0.62%		0.29%	0.97%	0.25%	0.19%	0.42%
	5	19.31%		0.70%	0.13%	1.17%	0.25%	0.21%	0.39%
AXP	4	13.21%	0.70%		0.79%	0.79%	0.19%	0.22%	0.42%
	5	13.79%		0.80%	0.32%	0.86%	0.21%	0.57%	0.46%
BA	4	11.33%	1.14%		1.53%	0.86%	1.62%	0.86%	0.53%
	5	11.27%		1.08%	0.84%	1.15%	1.57%	1.04%	0.53%
BAC	4	19.41%	0.95%		0.18%	0.39%	0.50%	0.34%	1.06%
	5	12.58%		0.60%	0.43%	0.39%	0.49%	0.42%	1.10%
CAT	4	14.19%	0.43%		0.13%	0.49%	0.65%	0.47%	0.39%
	5	13.77%		0.45%	0.07%	0.48%	0.58%	0.24%	0.39%
Stocks	4	12.19%	0.67%		0.57%	1.20%	0.58%	0.43%	0.52%
	5	11.97%		0.68%	0.37%	1.26%	0.59%	0.55%	0.53%

Equation 1.5, the mean $PR_{m,ALP}$ value is 0.68%, while the mean values of $PR_{i,r}$ and $PR_{i,\sigma}$ are 0.37% and 1.26%, respectively. This emphasises the need to account for the conditional nature of illiquidity persistence in illiquidity modelling.

1.5.5. Forecast Accuracy

The previous sub-section demonstrated the *ex post* explanatory power of ALP and the extended THAR(3) model. This sub-section shows the *ex ante* explanatory power of ALP and the extended THAR(3) model. The forecast performance of the extended THAR(3) model is compared to performances of the HAR(3) model and the AR(1) model. Accordingly, in this sub-section an out-of-sample forecast period from the 3rd of January 2012 to the 31st of December 2012 is considered.

Similar to the main analysis, it is necessary to control for seasonality during the forecast period. Accordingly, the modified Amihud (2002) values during the forecast period are detrended using the estimated coefficients derived from the Gallant et al. (1992) methodology and the 2001 to 2011

Table 1.7
Forecast Accuracy of the AR(1), HAR(3) and extended THAR(3) Models

This table presents the forecast accuracies of the AR(1), HAR(3) and extended THAR(3) models for the forecast period, the 3rd of January 2012 to the 31st of December 2012. The HAR(3) model is presented in Equation 1.2. The extended THAR(3) model for the market sample refers to Equation 1.3. The extended THAR(3) model for the stock sample refers to Equation 1.5. The threshold parameters for the extended THAR(3) models are taken from the respective estimations of the THAR(3) models using the pre-forecast period. Mean absolute forecast error and mean squared forecast error are denoted by MAFE and MSFE, respectively. For the stock sample, the number of times each model performed best is reported. There are 29 stocks in the forecast analysis, due to the delisting of Kraft Foods Inc in 2012.

	MAFE			MSFE		
	AR(1)	HAR(3)	THAR(3)	AR(1)	HAR(3)	THAR(3)
ASX	0.706	0.645	0.632	0.401	0.337	0.300
DOW	0.718	0.652	0.625	0.470	0.395	0.348
FTSE	1.000	0.908	0.882	0.587	0.441	0.389
NASD	0.465	0.420	0.399	0.316	0.280	0.254
NIKK	1.077	0.992	0.988	0.549	0.402	0.384
SP500	1.016	0.875	0.850	1.349	1.004	0.917
Stocks	0	10	19	0	9	20

main sample period. This ensures a robust analysis of *ex ante* forecast performance.

For each daily illiquidity observation in the forecast period, we estimate a one-day-ahead forecast using all available observations prior to the forecast day. For the market illiquidity THAR(3) forecasts, the threshold parameters ($A_{\{k,-\}}$ and $A_{\{k,+\}}$) are taken from the estimation of Equation 1.3 for the main sample period, 2001 to 2011, shown in Table 1.3. The stock illiquidity threshold parameters are taken from the estimation of Equation 1.5 for the same main sample period. Accordingly, "extended THAR(3) model" will refer to Equation 1.3 for the market index sample and Equation 1.5 for the stock sample for the remainder of the study. Thus, the threshold parameters are chosen *ex ante* of the forecast period. The delisting of Kraft Foods Inc in 2012 means that it is excluded from the forecast analysis.

To assess the forecast accuracies of the three models, two popular inverse measures of forecast accuracy are used: mean absolute forecast error (MAFE); and, mean squared forecast error (MSFE):

$$MAFE_j = \frac{1}{N_j} \sum_{n=1}^{N_j} |ILLIQ^D_{j,n} - \widehat{ILLIQ}^D_{j,n}|$$

$$MSFE_j = \frac{1}{N_j} \sum_{n=1}^{N_j} \left(ILLIQ^D_{j,n} - \widehat{ILLIQ}^D_{j,n} \right)^2$$

where $j = i, m$, N_j is the number of trading days in 2012, $ILLIQ^D_{j,n}$ is the true level of illiquidity and $\widehat{ILLIQ}^{D}_{j,n}$ is the forecasted level of illiquidity at day n, for stock i or market index m.

Table 1.7 presents the MAFE and MSFE values for the market illiquidity observations. The THAR(3) model outperforms the short-memory AR(1) model and HAR(3) model for all market indices. Further, Table 1.7 also demonstrates that the THAR(3) model has smaller MAFE and MSFE values for the majority of stocks. Thus, incorporating conditional persistence into illiquidity modelling significantly enhances *ex ante* estimation.

1.5.6. Robustness

One concern could be that ALP is not robust to the specification of our THAR(3) model. To address this concern, we split our lagged HAR(3) illiquidity variables into 10 bins, for each value of k. The heterogeneous lags, $\overline{ILLIQ}^{k}_{t-1}$, are split into these bins according to their contemporaneous lagged mean return value, $\bar{r}^{k}_{t-1}$. The bins are constructed to cover the range of observed $\bar{r}^{k}_{t-1}$ values. The bins are as such:

$$\begin{aligned}
\bar{r}^{D}_{t-1} \in &\{(-\infty,-3.2],(-3.2,-2.4],(-2.4,-1.6],(-1.6,-0.8],(-0.8,0],\\
&(0,0.8],(0.8,1.6],(1.6,2.4],(2.4,3.2],(3.2,\infty)\},\\
\bar{r}^{W}_{t-1} \in &\{(-\infty,-1.6],(-1.6,-1.2],(-1.2,-0.8],(-0.8,-0.4],(-0.4,0],\\
&(0,0.4],(0.4,0.8],(0.8,1.2],(1.2,1.6],(1.6,\infty)\},\\
\bar{r}^{M}_{t-1} \in &\{(-\infty,-0.8],(-0.8,-0.6],(-0.6,-0.4],(-0.4,-0.2],(-0.2,0],\\
&(0,0.2],(0.2,0.4],(0.4,0.6],(0.6,0.8],(0.8,\infty)\}.
\end{aligned}$$

To examine ALP, independent of our THAR(3) specification, we regress daily illiquidity on our lagged illiquidity bins and our heterogeneous lag return and volatility variables. The regression formula can be represented as such:

$$\begin{aligned}
ILLIQ^{D}_{t} = \alpha_0 &+ \beta_{-4,D}\mathbb{1}\{\bar{r}^{D}_{t-1} \leq -3.2\}\overline{ILLIQ}^{D}_{t-1} + \beta_{-4,W}\mathbb{1}\{\bar{r}^{W}_{t-1} \leq -1.6\}\overline{ILLIQ}^{W}_{t-1}\\
&+ \beta_{-4,M}\mathbb{1}\{\bar{r}^{M}_{t-1} \leq -0.8\}\overline{ILLIQ}^{M}_{t-1} + \sum_{\delta=-3}^{3}\Big[\beta_{\delta,D}\mathbb{1}\{0.8\delta < \bar{r}^{D}_{t-1} \leq 0.8(\delta+1)\}\overline{ILLIQ}^{D}_{t-1}\\
&+ \beta_{\delta,W}\mathbb{1}\{0.4\delta < \bar{r}^{W}_{t-1} \leq 0.4(\delta+1)\}\overline{ILLIQ}^{W}_{t-1}\\
&+ \beta_{\delta,M}\mathbb{1}\{0.2\delta < \bar{r}^{M}_{t-1} \leq 0.2(\delta+1)\}\overline{ILLIQ}^{M}_{t-1}\Big]\\
&+ \beta_{4,D}\mathbb{1}\{\bar{r}^{D}_{t-1} > 3.2\}\overline{ILLIQ}^{D}_{t-1} + \beta_{4,W}\mathbb{1}\{\bar{r}^{W}_{t-1} > 1.6\}\overline{ILLIQ}^{W}_{t-1}\\
&+ \beta_{4,M}\mathbb{1}\{\bar{r}^{M}_{t-1} > 0.8\}\overline{ILLIQ}^{M}_{t-1} + \sum_{k=D}^{M}\Big[\theta_k\bar{r}^{k}_{t-1} + \lambda_k\bar{\sigma}^{k}_{t-1}\Big] + \epsilon_t. \qquad (1.6)
\end{aligned}$$

The estimation results of Equation 1.6 and the market illiquidity sample are shown in Figure 1.4.

Panel A: Daily Asymmetric Illiquidity Persistence

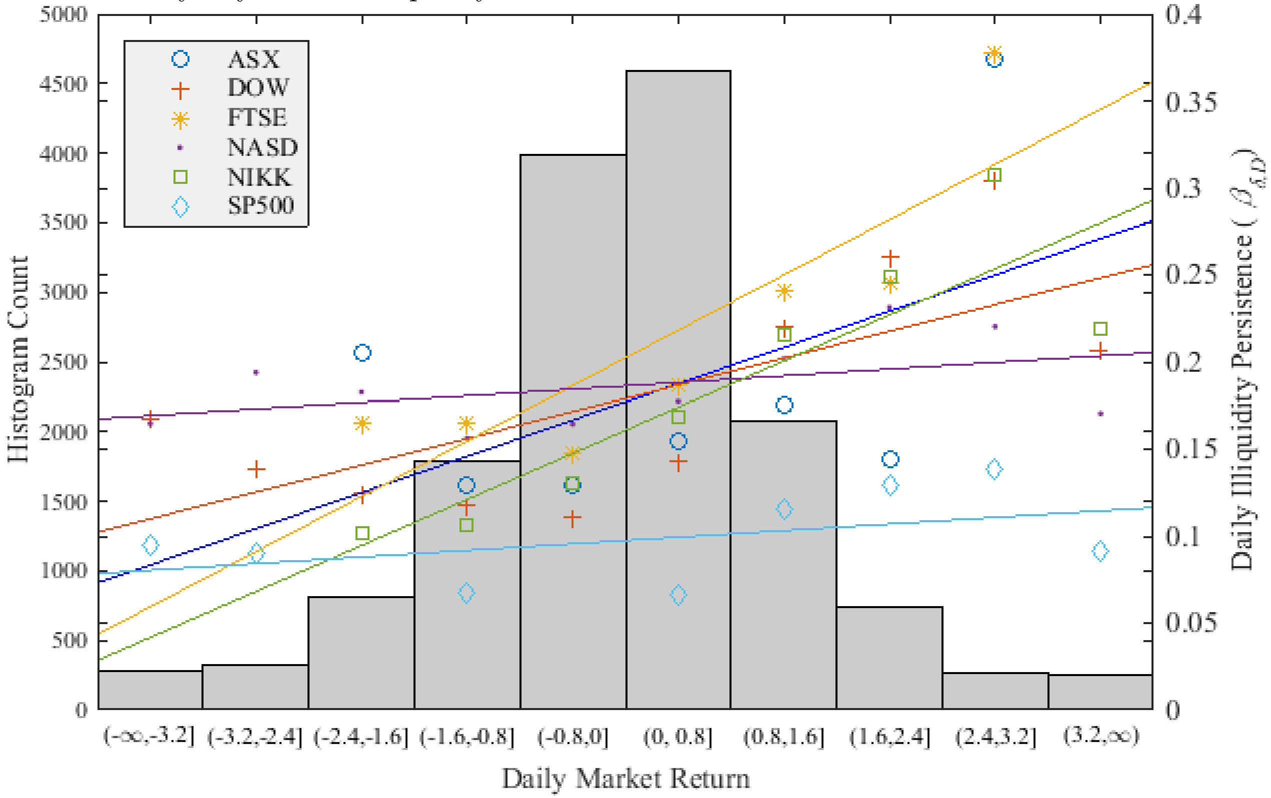
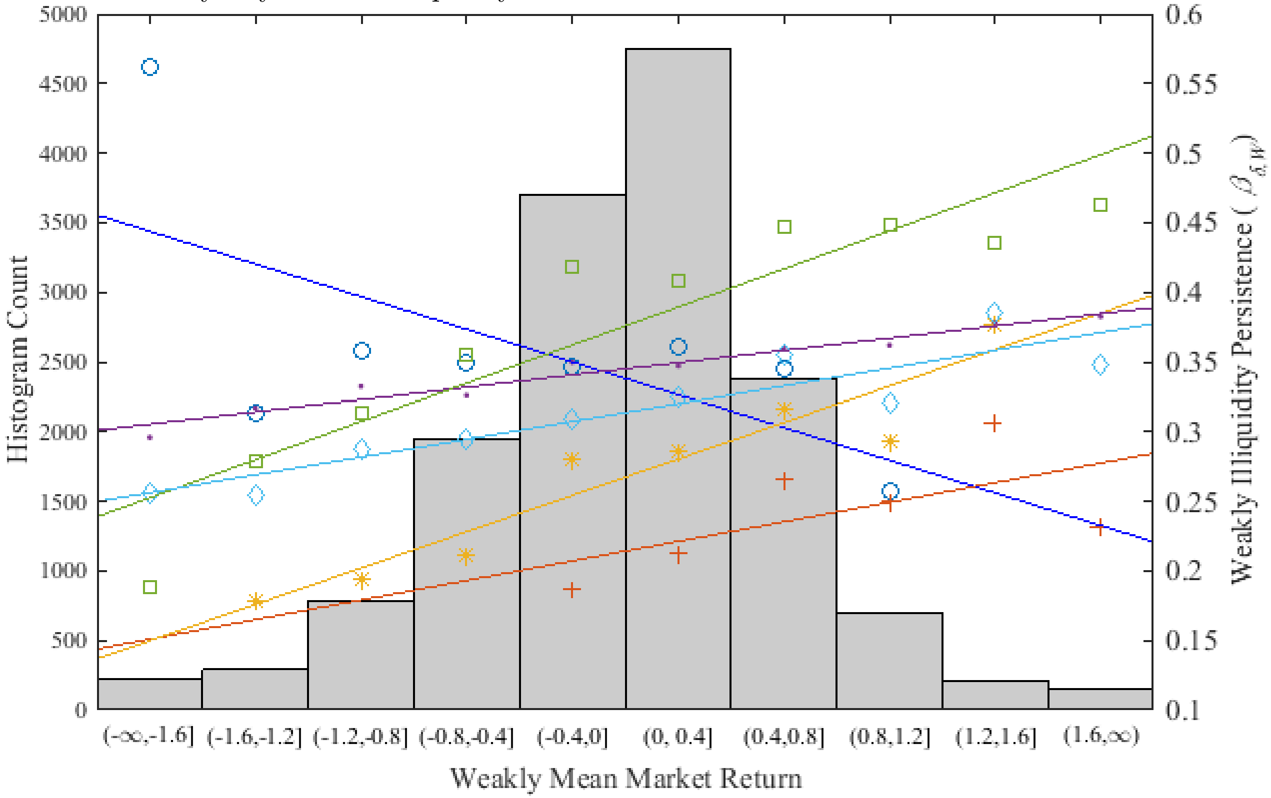


Fig. 1.4. Asymmetric Illiquidity Persistence. This figure shows the persistence of illiquidity as a function of lagged market returns. The figure is based on estimation of Equation 1.6:

$$\begin{aligned} ILLIQ_t^D = {} & \alpha_0 + \beta_{-4,D}\mathbb{1}\{\bar{r}_{t-1}^D \leq -3.2\}\overline{ILLIQ}_{t-1}^D + \beta_{-4,W}\mathbb{1}\{\bar{r}_{t-1}^W \leq -1.6\}\overline{ILLIQ}_{t-1}^W \\ & + \beta_{-4,M}\mathbb{1}\{\bar{r}_{t-1}^M \leq -0.8\}\overline{ILLIQ}_{t-1}^M + \sum_{\delta=-3}^{3}\Big[\beta_{\delta,D}\mathbb{1}\{0.8\delta < \bar{r}_{t-1}^D \leq 0.8(\delta+1)\}\overline{ILLIQ}_{t-1}^D \\ & + \beta_{\delta,W}\mathbb{1}\{0.4\delta < \bar{r}_{t-1}^W \leq 0.4(\delta+1)\}\overline{ILLIQ}_{t-1}^W + \beta_{\delta,M}\mathbb{1}\{0.2\delta < \bar{r}_{t-1}^M \leq 0.2(\delta+1)\}\overline{ILLIQ}_{t-1}^M\Big] \\ & + \beta_{4,D}\mathbb{1}\{\bar{r}_{t-1}^D > 3.2\}\overline{ILLIQ}_{t-1}^D + \beta_{4,W}\mathbb{1}\{\bar{r}_{t-1}^W > 1.6\}\overline{ILLIQ}_{t-1}^W \\ & + \beta_{4,M}\mathbb{1}\{\bar{r}_{t-1}^M > 0.8\}\overline{ILLIQ}_{t-1}^M + \sum_{k=D}^{M}\left[\theta_k\bar{r}_{t-1}^k + \lambda_k\bar{\sigma}_{t-1}^k\right] + \epsilon_t \end{aligned} \tag{1.6}$$

where $ILLIQ_t$ is the modified Amihud (2002) measure defined as $\ln\left(\frac{\sigma_t}{v_t}\right)$ where v_t is dollar volume at time t and σ_t is defined as $\ln\left(\frac{1}{2\sqrt{\ln 2}}\ln\left(\frac{P_t^H}{P_t^L}\right)\right)$, where P_t^H and P_t^L are the high and low prices on day t, respectively. Market returns are denoted by r_t. The independent variables are heterogeneous autoregressive lags corresponding to the $ILLIQ_t$, r_t and σ_t variables and superscript $k = D, W, M$, corresponding to "day", "week" and "month", respectively. Equation 1.6 is estimated for each market index within the sample. The coefficients, $\beta_{\delta,k}$ for $\delta \in \{-4,-3,-2,-1,0,1,2,3,4\}$, are displayed as a scatter plot, with trend line, over a combined histogram of $\bar{r}_{t-1}^k$, for all market indices. Panel A corresponds to $k = D$; Panel B to $k = W$ and Panel C to $k = M$. Only coefficients that are significant at the 10% level are shown, based on Newey-West robust standard errors.

Panel B: Weekly Asymmetric Illiquidity Persistence



Panel C: Monthly Asymmetric Illiquidity Persistence

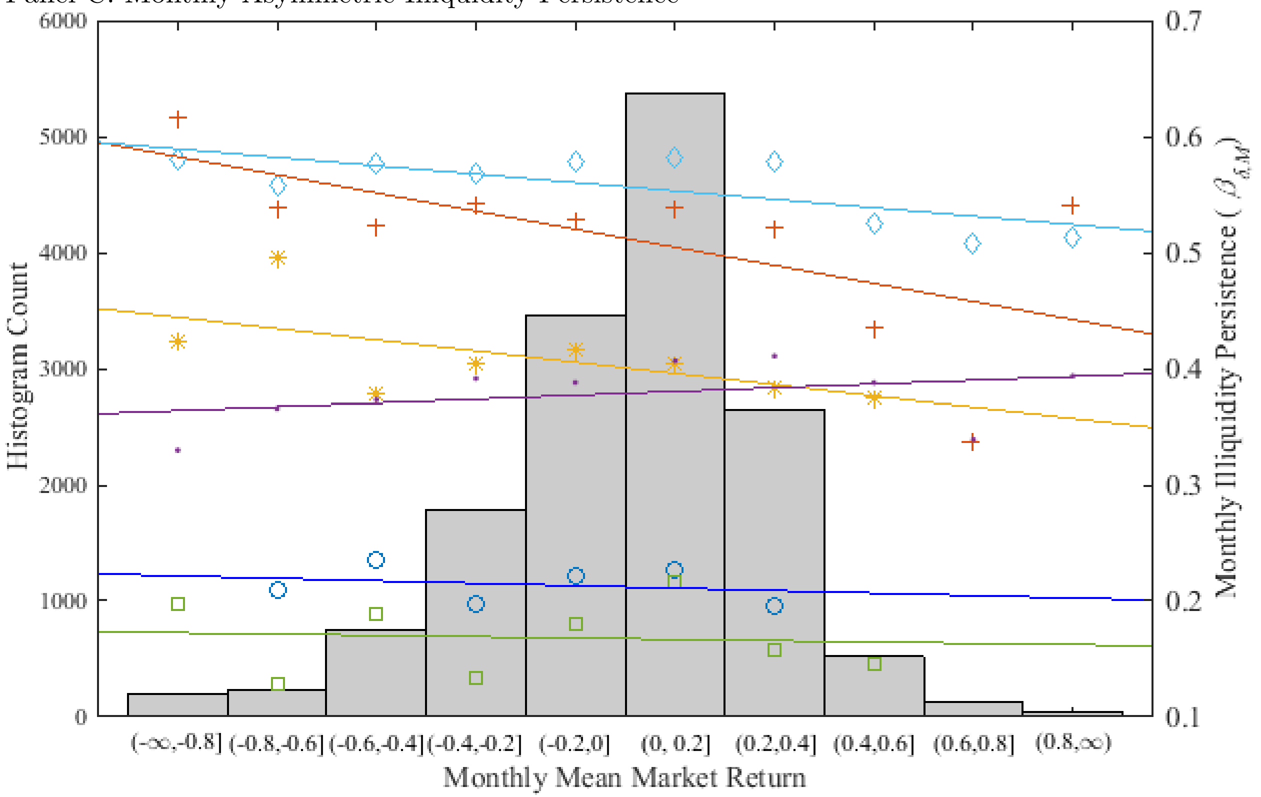


Fig. 1.4. (continued)

Panel A of Figure 1.4 plots the significant $\beta_{\delta,D}$ estimated coefficients for $\delta \in \{-4,-3,-2,-1,0,1,2,3,4\}$ from Equation 1.6. Panel A shows that all six markets display very strong positive ALP across the lagged daily return distributions. This is consistent with the conditional persistence values, captured by the THAR(3) model:

$$\hat{\rho}^{D}_{t-1}|(\mathscr{D}^{D}_{t-1}=-1) \leq \hat{\rho}^{D}_{t-1}|(\mathscr{D}^{D}_{t-1}=0) \leq \hat{\rho}^{D}_{t-1}|(\mathscr{D}^{D}_{t-1}=1).$$

Panel B plots the significant $\beta_{\delta,W}$ estimated coefficients for $\delta \in \{-4,-3,-2,-1,0,1,\ 2,3,4\}$. In this instance, the persistence of every illiquidity series is a positive function of past weekly returns, except that of the ASX, which exhibits a distinct negative weekly ALP. This reaffirms the results in Table 1.3 and Panel B of Figure 1.3, which document the ASX as the first market in the sample to revert to negative ALP, following a large positive return:

$$\hat{\rho}^{W}_{m,t-1}|(\mathscr{D}^{W}_{m,t-1}=-1) \leq \hat{\rho}^{W}_{m,t-1}|(\mathscr{D}^{W}_{m,t-1}=0) \leq \hat{\rho}^{W}_{m,t-1}|(\mathscr{D}^{W}_{m,t-1}=1) \quad \forall m \neq \text{ASX},$$

$$\hat{\rho}^{W}_{ASX,t-1}|(\mathscr{D}^{W}_{ASX,t-1}=-1) \geq \hat{\rho}^{W}_{ASX,t-1}|(\mathscr{D}^{W}_{ASX,t-1}=0) \geq \hat{\rho}^{W}_{ASX,t-1}|(\mathscr{D}^{W}_{ASX,t-1}=1).$$

Panel C demonstrates that at the monthly frequency, illiquidity persistence is mostly a negative function of past returns. This is again consistent with our estimation of the extended THAR(3) model given by Equation 1.3. Thus, it is apparent that illiquidity persistence is independent of our model specification and that our THAR model is able to capture ALP in a parsimonious and robust manner.

1.6. Amihud (2002) Hypotheses

In his influential paper, Amihud (2002) outlines two hypotheses regarding the time-series effects of illiquidity on stock returns. Amihud (2002) arrives at his hypotheses by assuming that illiquidity is persistent. This is confirmed in Table 1.1, where it is shown that daily illiquidity is not only persistence, but a long-memory process. Section 1.5 extends the analysis by demonstrating that the persistence of illiquidity is not constant, but conditional on past returns, which we call asymmetric liquidity persistence (ALP). Table 1.6 shows that ALP is a key determinant of daily illiquidity. Now, this section demonstrates the superior performance of the THAR(3) model, which accommodates ALP, in the context of the Amihud (2002) hypotheses.

Amihud (2002) begins his conjecture by identifying illiquidity as a transaction cost. He further assumes that investors value assets in terms of their net returns, rather than their raw returns. In the Amihud (2002) model, investors are assumed to predict the level of illiquidity for time t based on the information available at time $t-1$. Investors then use this information to set prices such that they can receive their desired expected return at time t. The expected return should compensate investors for the level of illiquidity experienced until time t. Therefore, expected stock returns are a positive function of expected illiquidity. Amihud's (2002) first hypothesis is:

H1: Ex ante stock returns are an increasing function of expected illiquidity.[7]

Amihud (2002) next assumes that illiquidity is persistent, such that unexpected illiquidity increases the expectations of illiquidity in the next time period. Thus, if prices are efficient,

H2: Unexpected illiquidity has a negative effect on contemporaneous unexpected stock returns.

This is Amihud's (2002) second hypothesis. This is analogous to Acharya and Pedersen's (2005) third proposition: "returns are low when illiquidity increases".

1.6.1. Illiquidity Innovations

In his paper, Amihud (2002) constructs his illiquidity innovations using an AR(1) specification. Hence, we begin by defining:

$$ILLIQ_t^D = \phi_0 + \phi_1 ILLIQ_{t-1}^D + ILLIQ_{AR(1),t}^{D,U}. \tag{1.7}$$

The residual from the AR(1) estimation of the illiquidity process is denoted by, $ILLIQ_{AR(1),t}^{D,U}$, where "U" signifies the "unexpected" component of illiquidity. Given that illiquidity has long-memory, the short-memory AR(1) model should be insufficient for capturing the persistence of $ILLIQ_t^D$, such that the residuals, $ILLIQ_{AR(1),t}^{D,U}$, are autocorrelated. On the other hand, the quasi-long-memory THAR(3) model should capture almost all of the persistence of daily illiquidity. To test this, we define $ILLIQ_{THAR(3),t}^{D,U}$ as the residual from Equation 1.3 for the market illiquidity series and Equation 1.5 for the stock illiquidity series. For completeness, we also define $ILLIQ_{HAR(3),t}^{D,U}$ as the residual series from a HAR(3) estimation, Equation 1.2.

Table 1.8 reports persistence statistics for the residual illiquidity series derived from the AR(1), HAR(3) and THAR(3) models. For all market $ILLIQ_{AR(1),t}^{D,U}$ series, the MRS statistic is greater than the 95% level of confidence critical value of 1.862. We reject the null hypotheses that the $ILLIQ_{AR(1),t}^{D,U}$ series do not have long-memory at the 95% confidence level. Thus, the AR(1) specification does not model the true dependence structure of illiquidity. In contrast, all market $ILLIQ_{THAR(3),t}^{D,U}$ series have an MRS statistic that is less than one. This indicates that the extended THAR(3) model successfully captures the dependence structure of illiquidity such that the residuals do not display any significant autocorrelation. Hence, the AR(1) specification of illiquidity is insufficient for modelling the persistence of illiquidity, while the extended THAR(3) specification is not.

[7] Acharya and Pedersen (2005) point out that if illiquidity is persistent, ex ante stock returns will also be an increasing function of current illiquidity. This forms the second proposition of their study.

Table 1.8
Persistence in "Unexpected" Illiquidity

This table reports the persistence in the residuals of the AR(1), HAR(3) and extended THAR(3) models of illiquidity. The AR(1) specification is given by Equation 1.7. The HAR(3) specification is given by Equation 1.2. The THAR(3) model refers to Equation 1.3 for the market illiquidity series and Equation 1.5 for the stock illiquidity series. ACF(1) is the first-order autocorrelation coefficient. Hurst is the Hurst (1951) exponent. MRS is Lo's (1991) MRS statistic. For consistency, for each index and stock, the number of lags needed to constructed the MRS statistic, is taken from the $ILLIQ^{D}_{j,t}$ series where $j = i, m$. The number of lags is then determined using the methodology of Lo (1991). Mean statistics are reported for the stock sample.

	$ILLIQ^{D,U}_{AR(1),t}$			$ILLIQ^{D,U}_{HAR(3),t}$			$ILLIQ^{D,U}_{THAR(3),t}$		
	ACF(1)	Hurst	MRS	ACF(1)	Hurst	MRS	ACF(1)	Hurst	MRS
ASX	-0.183	0.829	2.675	-0.012	0.530	-1.101	-0.016	0.531	-0.902
DOW	-0.238	0.839	2.729	0.001	0.519	0.772	-0.011	0.505	0.709
FTSE	-0.259	0.812	2.380	0.038	0.584	-0.093	0.028	0.583	-0.171
NASD	-0.218	0.810	2.188	-0.003	0.508	-0.926	-0.005	0.506	-0.806
NIKK	-0.220	0.826	2.811	0.059	0.678	-0.886	0.056	0.655	-0.857
SP500	-0.299	0.826	2.443	0.005	0.500	0.528	-0.007	0.493	0.611
Stocks	-0.172	0.809	2.464	0.044	0.613	-0.982	0.042	0.571	-0.915

1.6.2. Modelling and Testing the Amihud (2002) Hypotheses

To test the Amihud (2002) hypotheses, we adapt his empirical model to the daily frequency. Hence, we estimate:

$$r_t = \alpha_0 + \gamma r_{t-1} + \delta\sigma_{t-1} + \zeta_1 ILLIQ^{D}_{t-1} + \zeta_2 ILLIQ^{D,U}_{AR(1),t} + \epsilon_t. \tag{1.8}$$

Under Hypothesis H1, we should expect $\zeta_1 > 0$. Further, under Hypothesis H2, we should expect $\zeta_2 < 0$. The additional variables, r_{t-1} and σ_{t-1}, are used as control variables. Equation 1.8 is estimated for each market and stock within the sample.

Table 1.9 contains the estimated regression results for Equation 1.8. With respect to Hypothesis H1, the results are mixed. For the sample of market illiquidity, there is only one ζ_1 coefficient that is positive and significantly different from zero. That is, the estimated ζ_1 coefficient for NASD is 0.226 and significant at the 99% confidence level. For the sample of stock illiquidity, the mean estimated ζ_1 coefficient is not significantly greater than zero. The insignificant results with respect to Hypothesis H1 are in line with two recent papers that re-examine the Amihud (2002) hypotheses. Drienko, Smith, and von Reibnitz (2018) and Harris and Amato (2018) demonstrate that Amihud's (2002) first hypothesis does not hold for more recent sample periods.

The support for Amihud's (2002) second hypothesis is more compelling. Every estimated ζ_2 coefficient in Table 1.9 is negative and statistically significant. All estimated ζ_2 coefficients for the market sample are significantly different from zero at the 99% level of confidence. The mean estimated ζ_2 value for the stock sample is significantly different from zero at the 90% level of confidence. Thus, Amihud's (2002) second hypothesis holds at the daily frequency when using the

Table 1.9
Test of the Amihud (2002) Hypotheses Using AR(1) Illiquidity Innovations

This table reports the estimation results for the following regression:

$$r_t = \alpha_0 + \gamma r_{t-1} + \delta\sigma_{t-1} + \zeta_1 ILLIQ^D_{t-1} + \zeta_2 ILLIQ^{D,U}_{AR(1),t} + \epsilon_t \tag{1.8}$$

where r_t is the return of a stock or a market index on day t. The volatility measure, σ_t, is defined as $\sigma_t = \ln\left(\frac{1}{2\sqrt{\ln 2}}\ln\left(\frac{P^H_t}{P^L_t}\right)\right)$ where P^H_t and P^L_t are the high and low prices for a stock or market on day t, respectively. The illiquidity variable, $ILLIQ^D_t$, is defined as $ILLIQ^D_t = \ln\left(\frac{\sigma_t}{v_t}\right)$ where v_t is the dollar volume of a stock or market on day t. The illiquidity innovation, $ILLIQ^{D,U}_{AR(1),t}$, is the residual from the estimation of the AR(1) model defined by Equation 1.7. The t-statistics are reported in italics. The mean coefficients are reported for the stock sample, with their associated Hameed et al. (2010) t-statistics in italics. Newey-West robust standard errors are used throughout. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively.

	α_o	γ	δ	ζ_1	ζ_2	R^2
ASX	0.014	-0.065***	0.023	-0.005	-0.136***	0.010
	0.21	*-3.16*	*0.69*	*-0.12*	*-2.84*	
DOW	-0.009	-0.118***	0.040	-0.013	-0.133***	0.017
	-0.09	*-5.85*	*0.89*	*-0.31*	*-3.07*	
FTSE	0.055	-0.092***	0.030	0.021	-0.138***	0.016
	0.93	*-4.74*	*0.66*	*0.79*	*-4.24*	
NASD	0.572***	-0.109***	-0.172**	0.226***	-0.451***	0.028
	3.51	*-5.30*	*-2.32*	*3.60*	*-4.89*	
NIKK	0.024	-0.045**	0.040	0.007	-0.267***	0.022
	0.28	*-2.30*	*0.75*	*0.22*	*-5.55*	
SP500	0.035	-0.123***	0.043	0.004	-0.188***	0.023
	0.17	*-6.39*	*0.84*	*0.08*	*-3.69*	
Stocks	-0.015	-0.059***	-0.004	0.032	-0.089*	0.008
	-0.37	*-3.01*	*-0.09*	*0.87*	*-2.03*	

AR(1) illiquidity innovation specification.

Given the long-memory properties of illiquidity, it is unlikely that the AR(1) illiquidity specification can adequately decompose illiquidity into its expected and unexpected components. Therefore, it is necessary to test the Amihud (2002) hypotheses with respect to the illiquidity innovations given by the THAR(3) model. Hence, we extend the analysis by considering the equation:

$$r_t = \alpha_0 + \gamma r_{t-1} + \delta\sigma_{t-1} + \zeta_1 ILLIQ^{D,E}_{THAR(3),t} + \zeta_2 ILLIQ^{D,U}_{THAR(3),t} + \epsilon_t. \tag{1.9}$$

where $ILLIQ^{D,E}_{THAR(3),t}$ is the expected illiquidity level and $ILLIQ^{D,U}_{THAR(3),t}$ is the unexpected illiquidity level on day t given from the estimation of the extended THAR(3) model. Analogous to Equation 1.8, for the Amihud (2002) hypotheses to hold, we expect: $\zeta_1 > 0$; and, $\zeta_2 < 0$.

Table 1.10 presents the estimated regression results for Equation 1.9. Table 1.10 reveals that there is limited evidence in favour of Amihud's (2002) first hypothesis. The estimated ζ_1 coefficients for the market sample are not statistically significant, while the mean estimated coefficient for the

Table 1.10
Test of the Amihud (2002) Hypotheses Using THAR(3) Illiquidity Innovations

This table reports the estimation results for the following regression:

$$r_t = \alpha_0 + \gamma r_{t-1} + \delta\sigma_{t-1} + \zeta_1 ILLIQ^{D,E}_{THAR(3),t} + \zeta_2 ILLIQ^{D,U}_{THAR(3),t} + \epsilon_t \quad (1.9)$$

where r_t is the return of a stock or a market index on day t. The volatility measure, σ_t, is defined as $\sigma_t = \ln\left(\frac{1}{2\sqrt{\ln 2}}\ln\left(\frac{P^H_t}{P^L_t}\right)\right)$ where P^H_t and P^L_t are the high and low prices for a stock or market on day t, respectively. The expected illiquidity level, $ILLIQ^{D,E}_{THAR(3),t}$, is given from the estimation of the THAR(3) model defined by Equation 1.3 for the market sample and Equation 1.5 for the stock sample. The illiquidity innovation, $ILLIQ^{D,U}_{THAR(3),t}$, is the residual from the estimation of the THAR(3) model defined by Equation 1.3 for the market sample and Equation 1.5 for the stock sample. The t-statistics are reported in italics. The mean coefficients are reported for the stock sample, with their associated Hameed et al. (2010) t-statistics in italics. Newey-West robust standard errors are used throughout. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively.

	α_o	γ	δ	ζ_1	ζ_2	R^2
ASX	-0.119	-0.062***	0.039	-0.082	-0.127***	0.009
	-0.80	*-2.94*	*1.26*	*-1.02*	*-3.01*	
DOW	-0.061	-0.113***	0.035	-0.034	-0.142***	0.017
	-0.35	*-5.53*	*1.12*	*-0.50*	*-3.60*	
FTSE	0.086	-0.086***	0.008	0.040	-0.140***	0.016
	0.97	*-4.46*	*0.21*	*1.00*	*-4.41*	
NASD	0.381	-0.094***	-0.025	0.147	-0.619***	0.034
	1.23	*-4.51*	*-0.45*	*1.28*	*-6.67*	
NIKK	-0.025	-0.033*	0.046	-0.013	-0.252***	0.020
	-0.15	*-1.69*	*1.00*	*-0.21*	*-5.09*	
SP500	-0.131	-0.116***	0.054*	-0.039	-0.202***	0.022
	-0.47	*-6.06*	*1.68*	*-0.57*	*-4.06*	
Stocks	-0.055	-0.056***	-0.013	0.080	-0.132**	0.009
	-0.86	*-2.91*	*-0.28*	*1.16*	*-2.57*	

stock sample is also not significantly different from zero. Again, this result mirrors the findings of Drienko et al. (2018) and Harris and Amato (2018).

In contrast to the expected illiquidity component, the unexpected component of illiquidity demands a statistically significant negative coefficient. In Table 1.10, all estimated ζ_2 values for the market sample are negative and statistically significant at the 99% level of confidence. The mean estimated ζ_2 value for the stock sample is also negative and significant at the 95% level of confidence. Thus, Amihud's (2002) second hypothesis holds for both the AR(1) and THAR(3) illiquidity specifications.

1.6.3. Dynamics of the Amihud (2002) Hypotheses

The conjecture of Amihud (2002) can be extended to derive expectations regarding the cross-correlation function of returns and illiquidity shocks. Amihud (2002) assumes that investors form their expectations of future illiquidity based on all available information in the current time period. That is, if an investor is at time t, their expectation of the level of illiquidity at time $t+1$, $ILLIQ^{D,E}_{t+1}$, should be based on all available information at time t. Accordingly, unexpected illiquidity at time $t+1$, $ILLIQ^{D,U}_{t+1}$, should be orthogonal to $ILLIQ^{D,E}_{t+1}$ and $ILLIQ^{D,E}_{t+1}$ should be representative of all illiquidity shocks that occurred prior to $t+1$. Thus, if prices are efficient, we should expect all past illiquidity shocks to have no impact on current returns. That is,

$$\text{corr}\left(ILLIQ^{D,U}_{t+l}, r_t\right) = 0 \qquad \forall l < 0. \tag{1.10}$$

Furthermore, given Amihud's (2002) second hypothesis and the results observed in tables 1.9 and 1.10, we should expect a negative relationship between unexpected illiquidity and contemporaneous returns. That is,

$$\text{corr}\left(ILLIQ^{D,U}_{t}, r_t\right) < 0. \tag{1.11}$$

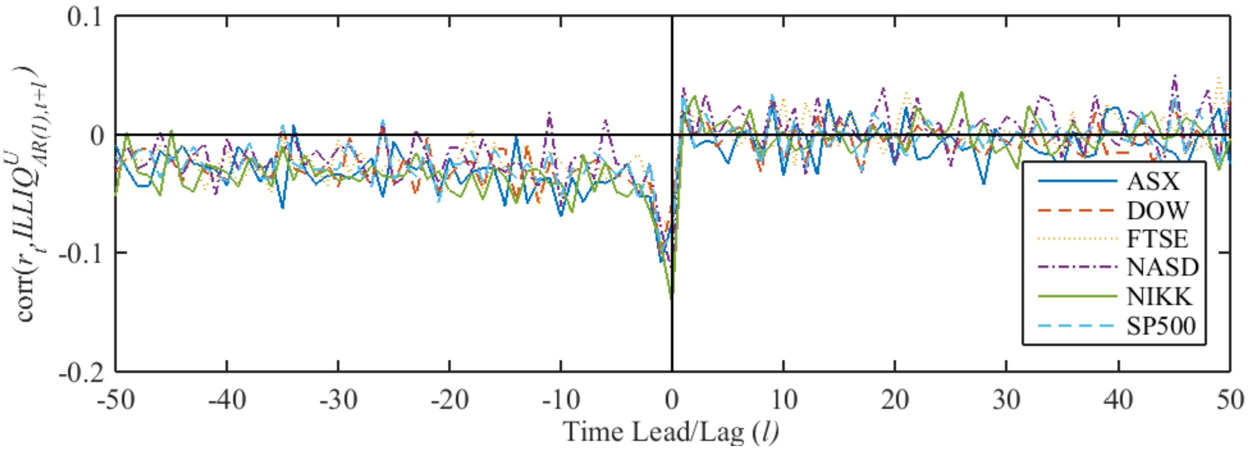
Additionally, unforeseeable future illiquidity shocks should not be priced. That is,

$$\text{corr}\left(ILLIQ^{D,U}_{t+l}, r_t\right) = 0 \qquad \forall l > 0. \tag{1.12}$$

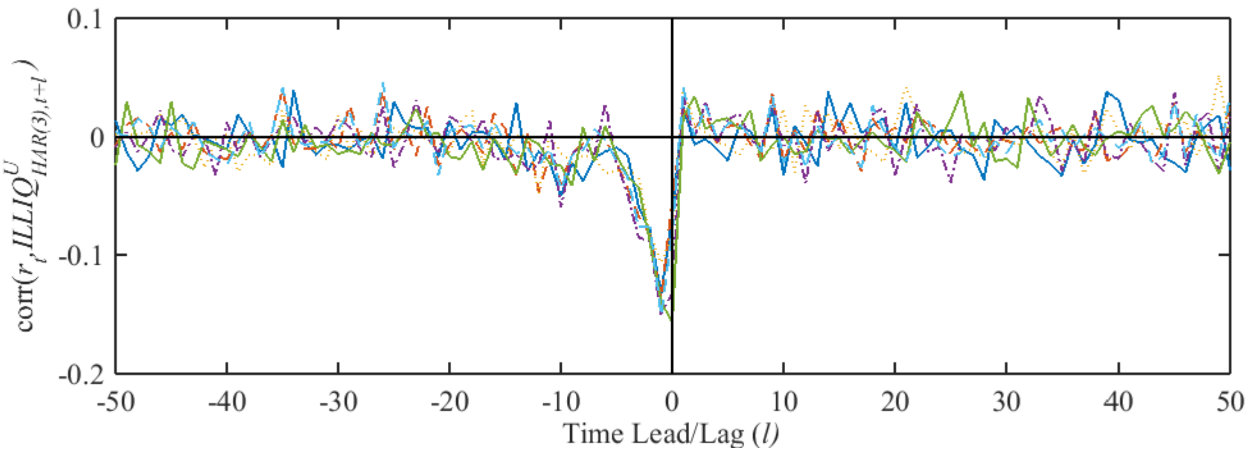
Figure 1.5 plots the cross-correlation functions of the returns and illiquidity innovations of the market indices in the sample. Panel A plots the cross-correlation functions for the AR(1) illiquidity innovations, $ILLIQ^{U}_{AR(1),t}$. Panel B plots the cross-correlation functions for the HAR(3) illiquidity innovations, $ILLIQ^{U}_{HAR(3),t}$. Panel C plots the cross correlation functions for the THAR(3) illiquidity innovations, $ILLIQ^{U}_{THAR(1),t}$. In each panel, we observe negative correlations between illiquidity residuals and contemporaneous returns. This is consistent with Amihud's (2002) second hypothesis and Inequality 1.11. Also consistent with Amihud's (2002) conjecture, future illiquidity shocks appear to be unrelated to contemporaneous returns. For $l > 0$, the cross-correlation functions are close to zero. Thus, the cross-correlation functions conform to Equation 1.12.

The variation across the panels of Figure 1.5 occurs in the $l < 0$ region. Panel A indicates that Equation 1.10 does not hold for the AR(1) model. In Panel A, the cross-correlation functions fail to converge to zero as l decreases towards -50. Thus, returns are negatively correlated to AR(1) illiquidity residuals from 50 days ago. This highlights a profound weakness of the AR(1) illiquidity specification. As shown in Table 1.8, the AR(1) illiquidity residuals exhibit long-memory. This causes returns to be negatively correlated to past illiquidity residuals. As suggested by the Amihud (2002) hypotheses, past unexpected illiquidity should be unrelated to current returns. Thus, Figure 1.5 demonstrates that the AR(1) specification does not adequately decompose illiquidity into its expected and unexpected components. The HAR(3) model improves on the AR(1) model. As evident in Panel B, the return-illiquidity residual cross-correlation functions converge to zero at approximately $l = -20$. The extended THAR(3) model has the best illiquidity innovation specification. Panel C of Figure 1.5 has the sharpest downward spike, with the cross-correlation functions converging to zero at approximately $l = -5$.

Panel A: Return and AR(1) Illiquidity Innovation Cross-Correlation Functions



Panel B: Return and HAR(3) Illiquidity Innovation Cross-Correlation Functions



Panel C: Return and THAR(3) Illiquidity Innovation Cross-Correlation Functions

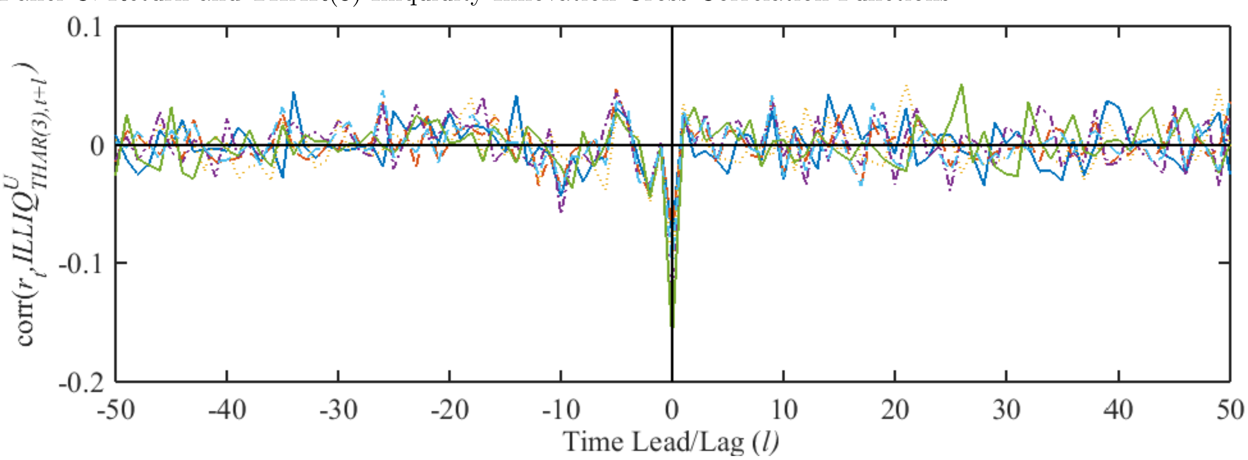


Fig. 1.5. Return-Illiquidity Residual Cross-Correlation Functions.
Panel A plots the cross-correlation functions for the daily return series and the AR(1) illiquidity innovations, corr $\left(r_t, ILLIQ^U_{AR(1),t+l}\right)$, for $l = -50, -49, ..., 50$. The AR(1) illiquidity innovations, $ILLIQ^{D,U}_{AR(1),t}$, are the residuals from the model defined by Equation 1.7. Panel B plots the equivalent functions for the HAR(3) illiquidity innovations, $ILLIQ^U_{HAR(3),t}$. The HAR(3) illiquidity innovations, $ILLIQ^{D,U}_{HAR(3),t}$, are the residuals from the model defined by Equation 1.2. Panel C plots the equivalent functions functions for the THAR(3) illiquidity innovations, $ILLIQ^U_{THAR(1),t}$. The THAR(3) illiquidity innovations, $ILLIQ^{D,U}_{THAR(3),t}$, are the residuals from the model defined by Equation 1.3.

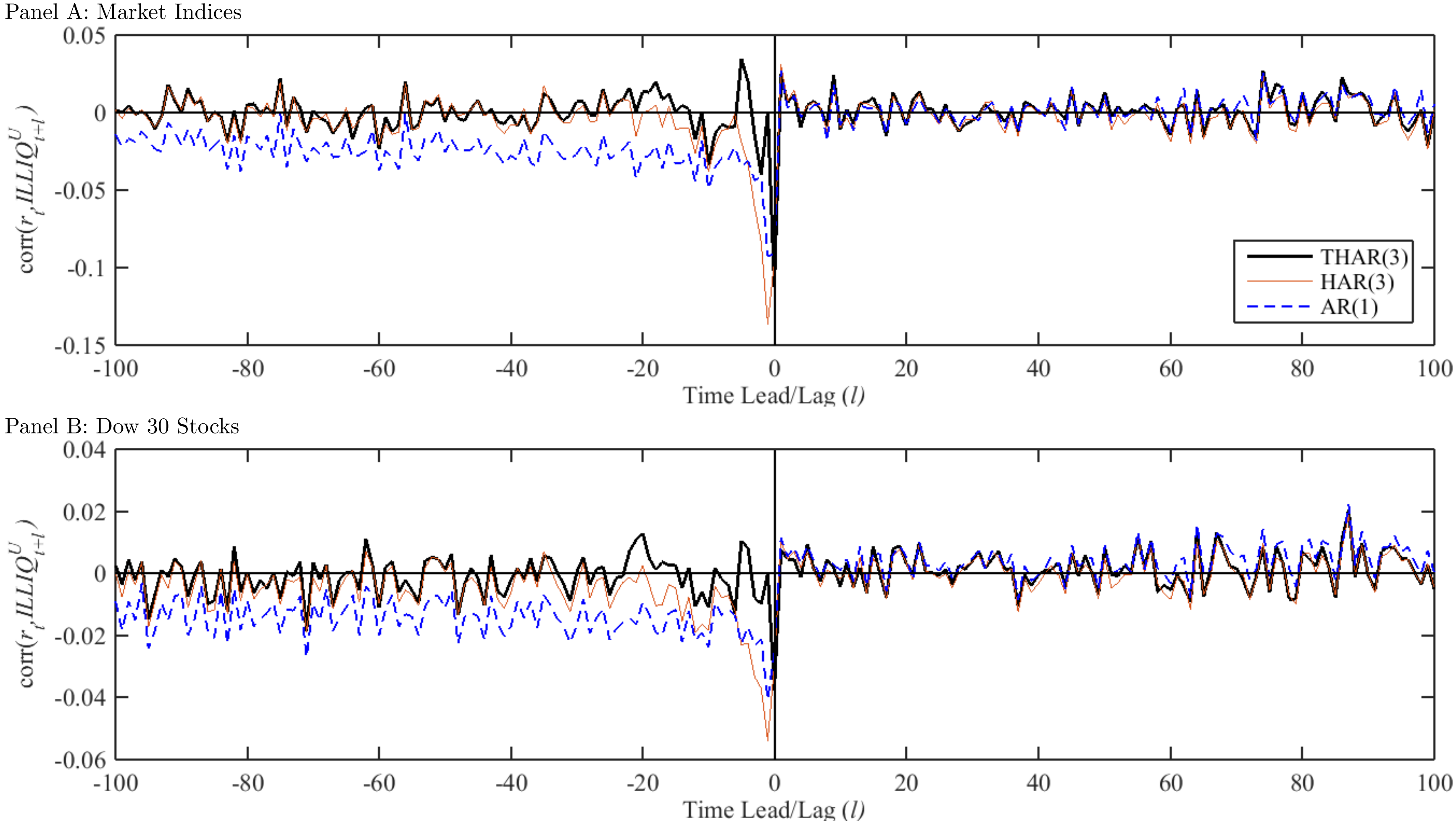


Fig. 1.6. Mean Return-Illiquidity Residual Cross-Correlation Functions.
Panel A plots the mean cross-correlation functions for the daily return series, r_t, and the market illiquidity innovations, $ILLIQ^U_{t+l}$, derived from the AR(1), HAR(3) and THAR(3) models. Panel B plots the mean cross-correlation functions for the daily return series and the stock illiquidity innovations derived from the AR(1), HAR(3) and THAR(3) models. The AR(1) illiquidity innovations, $ILLIQ^{D,U}_{AR(1),t}$, are the residuals from the model defined by Equation 1.7. The HAR(3) illiquidity innovations, $ILLIQ^{D,U}_{HAR(3),t}$, are the residuals from the model defined by Equation 1.2. The THAR(3) market illiquidity innovations, $ILLIQ^{D,U}_{THAR(3),t}$, are the residuals from the model defined by Equation 1.3. The THAR(3) stock illiquidity innovations are the residuals from the model defined by Equation 1.5.

Figure 1.6 plots the mean return-illiquidity residual cross-correlation functions for each illiquidity specification. Panel A plots the mean return-illiquidity innovation cross-correlation functions for the market indices in the sample. Panel B plots the mean return-illiquidity innovation cross-correlation functions for the stock sample. Similar to Figure 1.5, the mean return-illiquidity cross-correlation functions conform to Inequality 1.11 and Equation 1.12. Also in line with Figure 1.5, as l approaches -100 from above, the mean THAR(3) function converges to zero much faster than the HAR(3) and AR(1) specifications. Thus, according to expectations derived from the Amihud (2002) hypotheses, the THAR(3) model appears to provide for a superior illiquidity characterisation, for both market and stock illiquidity.

1.7. Conclusion

One well-known property of liquidity is that it is persistent. That is, current liquidity is strongly related to past levels of liquidity. Despite the persistence of liquidity, liquidity also has the tendency to rapidly deteriorate in market downturns. This paper reconciles these two contrasting phenomena by demonstrating that while liquidity is generally persistent, liquidity persistence is also sensitive to the state of the market.

This study begins by demonstrating that both daily market and stock liquidity have long-memory properties. In addition, we show that the persistence of liquidity is conditional on past stock market returns. This conditional liquidity persistence has very specific characteristics. Recent large negative returns can cause daily liquidity persistence to decrease by 90.7% (SP500); while, large negative returns in the more distant past increase liquidity persistence. This is indicative of a long-run equilibrium level of liquidity persistence. We term the conditional nature of liquidity persistence, asymmetric liquidity persistence (ALP).

We demonstrate the superiority of our conditional liquidity persistence model, the threshold heterogeneous autoregressive (THAR) model in three different contexts. First, we show that the in-sample explanatory power of ALP is comparable to the explanatory powers of both lagged returns and volatility. Second, we show that the THAR(3) model has greater out-of-sample forecasting accuracy than the short-memory AR(1) model and the unconditional persistence HAR(3) model. Finally, we illustrate the superior liquidity characterisation of the THAR(3) model under the theoretical predictions of Amihud (2002).

2. Discretionary Trading Surrounding Anticipated Distraction Events: the Case of the FIFA World Cup

2.1. Abstract

This study demonstrates how anticipated market-orthogonal events can induce discretionary trading. Following Ehrmann and Jansen (2017), this study uses FIFA World Cup matches that occur during trading hours as an exogenous shock to the opportunity cost of monitoring markets. World Cup football matches have an impact on contemporaneous trading and an asynchronous impact on the rest of the trading day. In particular, when World Cup matches occur in the middle of the trading day, there is an abnormally large amount of trading between market open and kick-off time. Dollar trading volume between 120 to 90 minutes before kick-off is 23.4% of a standard deviation higher than normal levels. This is due to a temporal substitution effect whereby traders submit their orders prior to kick-off in order to avoid trading during match time. During this pre-match period markets exhibit greater liquidity, volatility and price discovery. During matches, markets exhibit reduced liquidity, volatility and price discovery. The extraordinary market conditions that occur on match days follow the theoretical predictions of the Admati and Pfleiderer (1988) discretionary trading model.

JEL classification: G12, G14, G15, G41, L83.

Keywords: Limited Attention, Discretionary Trading, Intra-day Trading Patterns, FIFA World Cup.

2.2. Introduction

This study demonstrates how anticipated market-orthogonal events can induce discretionary trading. Following Ehrmann and Jansen (2017), this study uses FIFA World Cup matches that occur during trading hours as an exogenous shock to the opportunity cost of monitoring markets. The widespread appeal of the World Cup means that World Cup matches are an ideal device for observing exogenous shocks to the opportunity cost of monitoring markets. For example, 3.2 billion people watched at least one minute of the 2010 World Cup, while 909.6 million people watched at least one minute of the 2010 World Cup final.[8] In addition to the reduction in trading that occurs during and in the immediate vicinity of matches, as documented by Ehrmann and Jansen (2017), this paper demonstrates that World Cup football matches have a profound effect on the entire trading day. In particular, when World Cup matches occur in the middle of the trading day, there is an abnormal positive amount of trading in the pre-match trading period. In particular, trades between 120 to 90 minutes before kick-off are 21.3% of a standard deviation higher than normal

[8] 2010 FIFA World Cup South Africa: Television Audience Report. http://www.fifa.com/mm/document/affederation/tv/01/47/32/73/2010fifaworldcupsouthafricatvaudiencereport.pdf.

levels.

The pre-match period is also characterised by exceptional market conditions. Price impact costs are reduced by 14.2% of a standard deviation between 150 to 120 minutes before kick-off time. For the period 120 to 90 minutes before kick-off, volatility is increased by 21.5% of a standard deviation. These market conditions cannot be explained by the documented seasonal or intra-day trends of market activity. Further, the pre-match market conditions cannot be explained by the contemporaneous distraction effect of World Cup football matches, as described by Ehrmann and Jansen (2017).

The abnormal trading that occurs during the pre-match period is consistent with the Admati and Pfleiderer (1988) notion of intra-day discretionary trading. In the Admati and Pfleiderer (1988) model, discretionary liquidity traders concentrate their trades into one period of increased liquidity to reduce their transaction costs. This paper argues that discretionary traders not only consider their transaction costs but also the opportunity cost of monitoring the market. The opportunity cost of monitoring the market includes the cost of missing culturally significant events such as World Cup football matches. In some cases, the opportunity cost of monitoring the market can be great enough to induce traders to exit the market. For the case of World Cup football matches that occur during trading hours, discretionary traders with an interest in football are incentivised to fulfil their trading demand prior to kick-off time to ensure that they are able to dedicate their attention to monitoring the upcoming football match. Furthermore, according to the pooling Nash equilibria of Admati and Pfleiderer (1988), it is also optimal for discretionary traders not interested in football to trade in the pre-match period. This is because the pre-match period allows for the greatest number of market participants to pool their trades, football fans and non-football fans alike. This observation demonstrates how relatively minor market-orthogonal events can significantly affect financial markets.

This paper contributes to three distinct areas of the literature. First, this paper contributes to the growing literature that examines the impacts of limited investor attention on financial markets. This literature shows that individual investors are more likely to purchase stocks that grab their attention (Barber and Odean (2008); Seasholes and Wu (2007)), under-react to earnings announcements when they are distracted (Dellavigna and Pollet (2009); Hirshleifer, Lim, and Teoh (2009)) and trade less when they are distracted (Ehrmann and Jansen (2017)).

This study documents a new channel through which limited investor attention impacts on financial markets. Investor attention can influence discretionary trading behaviour. Accordingly, market-orthogonal events can have an asynchronous discretionary trading effect, as well as a contemporaneous distraction effect. This paper shows that the discretionary trading effect has important implications for the price impact costs of traders, market volatility and price discovery. Moreover, the incentives for discretionary traders to pool their trades together mean that even minor distraction events can have a significant impact on markets.

Second, this paper contributes to the growing literature that examines the impacts of sporting events on financial markets. The literature shows that there are two distinct channels through which sporting events impact financial markets. The first channel is through investor sentiment. The sports sentiment literature argues that sporting outcomes impact investor moods, which in turn affects investors' levels of optimism and pessimism. These shifts in optimism influence the buying and selling behaviour of investors. The sports sentiment effect has been documented by Ashton, Gerrard, and Hudson (2003), Edmans et al. (2007), Mishra and Smyth (2010) and Chang, Chen, Chou, and Lin (2012). In addition to the sports sentiment effect, Pantzalis and Park (2014) show that local sports sentiment is a determinant of local stock comovement. Kaplanski and Levy (2010) propose a trading strategy for exploiting the sports sentiment effects of FIFA World Cups. Their follow up paper, Kaplanski and Levy (2014), suggests that sophisticated investors did indeed exploit sentiment effects during the 2014 World Cup.

In addition to influencing investors' mood, sporting events can also increase the opportunity cost of monitoring the market. There are few studies in this area. This could be because most sporting events take place in the evening and outside of trading hours. Nonetheless, Wang and Markellos (2015) demonstrate that most countries experience a decline in trading activity during the Olympic Games. Ehrmann and Jansen (2017) recognise that during FIFA World Cup tournaments, there are occasions in which national football teams play matches within their domestic stock exchange trading hours. In particular, Ehrmann and Jansen (2017) point out that the 2010 FIFA World Cup featured 21 matches in which one of the competing nations' national stock exchanges was simultaneously open for trading. Ehrmann and Jansen (2017) argue that limited investor attention coupled with the distraction of FIFA World Cup matches caused trades to drop by an average of 38.0% and trading volume by 35.8%, during 2010 World Cup matches. This paper indicates that this was not exactly the case. For anticipated market-orthogonal events, there is a distraction effect *and* a discretionary trading effect. In the case of World Cup matches that occur during trading hours, discretionary traders substitute trading during match time for trading before kick-off time. Thus, this paper reveals a behavioural reaction to limited attention that has not previously been documented in the literature.

Third, this paper contributes to the empirical market microstructure literature that tests the predictions of the Admati and Pfleiderer (1988) model of discretionary trading. The Admati and Pfleiderer (1988) model predicts that during increased discretionary trading, volatility is increased, price discovery is increased and price impact is reduced. Further, when there is decreased discretionary trading, volatility is reduced, price discovery is reduced and price impact costs are increased. The empirical evidence regarding the theoretical predictions of the Admati and Pfleiderer (1988) model is mixed. For example, Foster and Viswanathan (1993) find evidence against Admati and Pfleiderer (1988), while the empirical results of Brailsford (1996), Scalia (1998) and Chae (2005) support Admati and Pfleiderer (1988).

This paper differs from the previous empirical tests of the Admati and Pfleiderer (1988) model by identifying an explicit discretionary trading preference. In the existing literature, the

Admati and Pfleiderer (1988) predictions are tested against variation in trading volumes over indiscriminate periods of time.[9] This paper shows that the Admati and Pfleiderer (1988) model predictions hold when discretionary trading is properly identified. Further, the Admati and Pfleiderer (1988) predictions hold during both increased discretionary trading and decreased discretionary trading. When there is increased discretionary trading during the pre-match period of match days, markets have lower price impact costs, higher volatility and higher price discovery. When there is decreased discretionary trading during match times, markets experience greater price impact costs, reduced volatility and reduced price discovery.

The remainder of this paper is structured as follows. Section 2.3 describes the data and provides summary statistics. Section 2.4 provides evidence of the intra-day discretionary trading that takes place on match days. Section 2.5 examines the implications of the intra-day discretionary trading for liquidity, volatility and price discovery. In Section 2.6, the findings of sections 2.4 and 2.5 are set against various robustness tests. The final section, Section 2.7, summarises the key findings of the paper.

2.3. Data and Summary Statistics

2.3.1. Football Match Data

The FIFA World Cup is a football competition that is organised by the Fédération Internationale de Football Association (FIFA). The tournament is held every four years. In each edition of the tournament, 32 senior men's national football teams compete for the World Cup Trophy and the title of 'World Champion'.[10] The World Cup is typically hosted in one country. This means that, as countries congregate in one time zone to participate in football matches within the host country's afternoon and evening hours, matches can be scheduled during the domestic trading hours of the participating teams in a quasi-random fashion. Due to the availability intra-day data, this study considers all World Cup football matches from 1998.

Table 2.1 gives the countries that have participated in a World Cup match during their own domestic trading hours since 1998, conditional on intra-day stock market data availability. In total, their are 98 football matches that occurred during trading hours from 1998 to 2014. The 98 match observations involve 22 countries. Table 2.1 also demonstrates that the 2010 World Cup featured the most matches during domestic trading hours. The 2010 World cup featured 29 matches during trading hours. Further, the 2010 World Cup features the largest cross-section of countries that participated in football matches during trading hours. The 29 matches during the 2010 World Cup that occurred during trading hours encompass 16 participating countries.

[9] The one exception is the Chae (2005) study. Chae (2005) finds evidence in favour of Admati and Pfleiderer (1988) but does not test all the predictions of the Admati and Pfleiderer (1988) model.

[10] There is also a World Cup that is contested by senior women's national football teams that is referred to as the 'FIFA Women's World Cup'.

Table 2.1
World Cup Football Matches during Trading Hours

This table indicates the number of times each country was involved in a FIFA World Cup match during their own domestic trading hours, for each World Cup since 1998 and conditional on intra-day stock market data availability.

Country	1998	2002	2006	2010	2014	1998 - 2014
Argentina	2		2	1	2	7
Belgium	1	3				4
Brazil	3		3	4		10
Chile				2	2	4
Colombia					3	3
Denmark	1			1		2
England	1	1		1		3
France	2	2		1		5
Germany	1	4	2	1		8
Greece				1		1
Ireland		2				2
Italy			2	1		3
Mexico	2		2	3	2	9
Netherlands	1			3		4
Poland		3	1			4
Portugal		2	1	3		6
Russia		2				2
South Africa				2		2
Spain		2	2	1		5
Switzerland			1	2		3
Turkey		4				4
United States	2		2	2	1	7
Total	16	25	18	29	10	98

The football match data is extracted from the official FIFA match reports available from the FIFA website. Each match report gives the scheduled date and time of kick-off, the final match outcome, the half-time score, the location of the match, the minute in which each goal was scored (for example, 1 to 90 if there is no extra stoppage time nor extra time) and the amount of extra stoppage time for each half. This information allows me to infer when each match was played and when each goal was scored with an accuracy of one minute, conditional on matches starting on time and half-time being exactly 15 minutes.

2.3.2. Market Data

The market data is comprised of five sub-samples. Each sub-sample corresponds to an edition of the World Cup. Since every World Cup is predominantly held in the month of June, I extract stock market data from the months of May, June and July of each World Cup year. This is to include a significant amount of time before and after each World Cup. Following

Table 2.2
Stock Market Trading Hours
This table details the trading hours of each national stock market during each iteration of the FIFA World cup from 1998 to 2014.

Country	1998	2002	2006	2010	2014
Argentina	11:00-17:00		11:00-17:00	11:00-17:00	11:00-17:00
Belgium	10:00-16:30	9:00-17:20			
Brazil	10:00-17:00		10:00-17:00	10:00-17:00	
Chile				9:30-16:00	9:30-16:00
Colombia					8:30-15:00
Denmark	9:00-17:00			9:00-17:00	
England	8:30-16:30 (1/5-19/7) 9:00-16:30 (20/7-31/7)	8:00-16:30		8:00-16:30	
France	10:00-17:00	9:00-17:30		9:00-17:30	
Germany	8:30-17:00	9:00-20:00	9:00-17:30	9:00-17:30	
Greece				10:30-17:20	
Ireland		8:00-17:30			
Italy			9:00-17:25	9:00-17:25	
Mexico	8:30-15:00		8:30-15:00	8:30-15:00	8:30-15:00
Netherlands	9:30-16:30			9:00-17:30	
Poland		10:00-16:00	10:00-16:00		
Portugal		8:00-16:30	8:00-16:30	8:00-16:30	
Russia		10:00-18:00			
South Africa				9:30-17:00	
Spain		9:00-17:30	9:00-17:30	9:00-17:30	
Switzerland			9:00-17:20	9:00-17:20	
Turkey		9:30-12:00 & 14:00-16:30			
United States	9:30-16:00		9:30-16:00	9:30-16:00	9:30-16:00

Ehrmann and Jansen (2017), the cross-section of each World cup sub-sample is determined by two criteria. First, the country must have intra-day stock market data available. Second, the country's national stock exchange must be open for trading during at least one match in which the country's national football team is participating. The trading hours of each country during each World Cup are presented in Table 2.2.

There are 22 countries in the sample: 15 European, six American and South Africa. I take the constituent stocks of each country's national index to construct my stock sample. I identify the constituent stocks using Thomson Reuters Tick History (TRTH). For each stock, I use five-minute intra-day data, and trade and quote data from TRTH.[11] [12] I conduct my analysis at the five-minute frequency. This is to mitigate any potential measurement error with regards to the football match data. For example, football matches often start a couple of minutes behind schedule and half-time is not always exactly 15 minutes.

[11]I gratefully acknowledge the Securities Industry Research Centre of Asia-Pacific (SIRCA) for making this data available.
[12]For a detailed explanation of the TRTH database, see Fong, Holden, and Trzcinka (2017).

Table 2.3
Market Indices
This table details the market index sample. The representative stocks for each country are sampled from the following national indices.

Country	1998	2002	2006	2010	2014
Argentina	MERV		MERV	MERV	MERV
Belgium	BFX	BFX			
Brazil	BVSP		BVSP	BVSP	
Chile				IPSA	IPSA
Colombia					COLCAP
Denmark	KFX			OMXC20	
England	FTSE	FTSE		FTSE	
France	FCHI	FCHI		FCHI	
Germany	GDAXI	GDAXI	GDAXI	GDAXI	
Greece				ATF	
Ireland		ISEQ			
Italy			SPMIB	FTMIB	
Mexico	MXX		MXX	MXX	MXX
Netherlands	AEX			AEX	
Poland		WIG20	WIG20		
Portugal		PSI20	PSI20	PSI20	
Russia		IRTS			
South Africa				JDTOP	
Spain		IBEX	IBEX	IBEX	
Switzerland			SSMI	SSMI	
Turkey		XU030			
United States	DJI		DJI	DJI	DJI

Table 2.3 identifies the market indices used to arrive at each market sub-sample. The market indices are identified by their Reuters Instrument Codes. The market index constituents are not constant over the sub-sample time periods. To ensure that each market sub-sample has a constant number of stocks, each market sub-sample includes all stocks that were in the corresponding market index during the sub-sample period. For example, London Stock Exchange Group and Thomas Cook Group were replaced by African Barrick Gold and Essar Energy in the FTSE 100 on the 21st of June 2010. Accordingly, London Stock Exchange Group, Thomas Cook Group, African Barrick Gold and Essar Energy are all included in the 2010 England sub-sample.

2.3.3. Trading Activity Variables

To observe trading activity, I construct 11 market variables. Market volume, $VOL_{m,t,w}$, is calculated by aggregating the trading volumes of the constituent stocks of market index, m, at time t, in World Cup sub-sample w. Market dollar volume ($DVOL_{m,t,w}$), number of trades ($TRADES_{m,t,w}$), number of bids at the national best bid price ($NBBOBIDS_{m,t,w}$), number of asks at the national best ask price ($NBBOASKS_{m,t,w}$), total number of bids ($BIDS_{m,t,w}$) and total number of asks ($ASKS_{m,t,w}$) are constructed in a similar manner. The $BIDS_{m,t,w}$ and $ASKS_{m,t,w}$ variables are only available for the USA and the Euronext countries in the sample. I also classify trades into buyer-initiated and seller-initiated trades using the Lee and Ready (1991) algorithm. The amount of buyer-initiated dollar trading volume and seller-initiated dollar trading volume for market m at time t in World Cup sub-sample w are given by $BDVOL_{m,t,w}$ and $SDVOL_{m,t,w}$, respectively.

Another important form of market activity in order-driven markets is limit order revisions and cancellations. Not all limit order revisions nor cancellations can be directly observed from trade and quote data. For example, if there is an increase in the bid price, it is unclear if this is due to a new bid order or a revised bid order. On the other hand, if there is a decrease in the bid price that does not immediately follow a trade, this can be identified as a bid revision or cancellation. Similarly, if a quote size decreases without a trade taking places, a limit order revision can be inferred. Accordingly, it is possible to approximate for bid and ask revisions using trade and quote data. The bid order revision proxy, $BREVISIONS_{m,t,w}$, counts the number of times the best bid price or volume at the best bid price for stocks in market m, at time t, in World Cup sub-sample w decrease without a trade taking place. The ask order revision proxy, $AREVISIONS_{m,t,w}$ counts the number of times the best ask price increases or volume of the best ask price decreases for stocks in market m, at time t, in World Cup sub-sample w decrease without a trade taking place.

2.3.4. Market Condition Variables

This section outlines the methodology for constructing the price impact, volatility and price discovery measures used in this study. I use two measures of price impact in the main analysis. The first is the Amihud (2002) measure. The Amihud (2002) measure, $AMIHUD_{i,t,m}$, is defined as:

$$AMIHUD_{i,t,m} = \frac{|r_{i,t,m}|}{DVOL_{i,t,m}}.$$

where $r_{i,t,m}$ and $DVOL_{i,t,m}$ are the continuously compounded return and dollar trading volume for stock i, at time t in World Cup sub-sample w, respectively. The Amihud (2002) ratio is a price impact proxy used throughout the market microstructure and asset-pricing literatures.[13]

For extremely illiquid periods in which no trades are recorded, $AMIHUD_{i,t,w}$ cannot be calculated. For these periods denoted by t', I let $AMIHUD_{i,t',w} = \max(AMIHUD_{i,t,w})$. That is, I equate $AMIHUD_{i,t',w}$ to the most illiquid observation available for stock i in World Cup sub-sample w. Following, I calculate a market level Amihud (2002) measure, $AMIHUD_{m,t,w}$, by taking the mean of $AMIHUD_{i,t,w}$ across market index constituents at time t in World Cup sub-sample w.

I consider a second price impact measure for robustness. The second price impact measure is a modified version of the Amihud (2002) measure. The modified Amihud (2002) measure is used

[13] Lou and Shu (2017) report that from 2009 to 2015 over 120 papers published in the Journal of Finance, the Journal of Financial Economics and the Review of Financial Studies use the Amihud (2002) measure in their empirical analysis.

in Chapter 1, Asymmetric Liquidity Persistence, to overcome some of the documented weaknesses of the Amihud (2002) measure. For example, Andersen and Bollerslev (1998) show that absolute returns are a poor measure of price movement; while, Lou and Shu (2017) show that most of the time-series variation in the Amihud (2002) measure can be attributed to the denominator. The modified Amihud (2002) measure for stock i at time t in World Cup sub-sample w is defined as:

$$MAMIHUD_{i,t,w} = \frac{\sigma_{i,t,w}}{DVOL_{i,t,w}}$$

where $\sigma_{i,t,w}$ is a volatility measure that utilises the five-minute high-price, $P^H_{i,t,w}$, and five-minute low price, $P^L_{i,t,w}$,:

$$\sigma_{i,t,w} = \frac{1}{2\sqrt{\ln 2}} \ln\left(\frac{P^H_{i,t,w}}{P^L_{i,t,w}}\right).$$

Similar to the $AMIHUD_{i,t,w}$ measure, for periods in which there are zero trades, denoted by t', I let $MAMIHUD_{i,t',w} = \max\left(MAMIHUD_{i,t,w}\right)$.

The market level modified Amihud (2002) measure, $MAMIHUD_{m,t,w}$, is taken as the cross-sectional average $MAMIHUD_{i,t,w}$ value of the market index constituents at time t. Further, the market level volatility measure, $\sigma_{m,t,w}$, is taken as the cross-sectional average $\sigma_{i,t,w}$ value of the market index constituents at time t.

In addition to the Amihud (2002) and modified Amihud (2002) measures, I also consider the high-frequency liquidity measures such as the effective spread, the realised spread and five-minute price impact. These measures are discussed and analysed in Section 5.2.4 in Chapter 5 Limitations and Future Research.

The methodology for calculating the market-level price discovery measure largely follows from Wang and Yang (2017). The Wang and Yang (2017) study utilises the Beveridge and Nelson (1981) decomposition to separate five-minute returns into its permanent and transitory components. Following Wang and Yang (2017), I begin by estimating the following equation for each stock on a daily basis using ordinary least squares:

$$r_{i,t,d,w} = A(L)_{i,d,w} r_{i,t,d,w} + u_{i,t,d,w}$$

where $r_{i,t,d,w}$ is the continuously compounded return of stock i at five-minute time period t on day d in World Cup sub-sample w, $A(L)$ is a lag operator and $u_{i,t,d,w}$ is the error term. The lag operator is defined as:

$$A(L)_{i,d,w} = \alpha_{i,d,1,w} L + \alpha_{i,d,2,w} L^2 + \alpha_{i,d,3,w} L^3 + \alpha_{i,d,4,w} L^4 + \alpha_{i,d,5,w} L^5 + \alpha_{i,d,6,w} L^6.$$

For simplicity, I keep the number of lags constant at six. Wang and Yang (2017) allow for up to six lags and determine the number of lags by taking the average of the Akaike (1974) and Schwarz (1978) Bayesian information criteria. According to Wang and Yang (2017), the permanent return component for stock i at five-minute time period t, day d, in World Cup sub-sample w can be estimated as:

$$\Delta m_{i,t,d,w} \equiv \left(\frac{\hat{u}_{i,t,d,w}}{(1 - A_{i,d,w})} \right)$$

where $A_{i,d,w}$ is:

$$A_{i,d,w} = \sum_{n=1}^{6} \hat{\alpha}_{i,d,n,w}.$$

In the spirit of Hasbrouck (1991), the variance of $\Delta m_{i,t,d,w}$ can be interpreted as trade informativeness. Accordingly, Wang and Yang (2017) arrive at their daily gross price discovery measure for stock i and day d by taking the sum of $(\Delta m_{i,t,d,w})^2$ over each intraday sub-period t. Since I am interested in constructing an intra-day gross price discovery measure, I take a different approach. I estimate my market level gross price discovery measure, $PD_{m,t,w}$, by taking the logarithmic function of the cross-sectional mean of the $(\Delta m_{i,t,d,w})^2$ values of the market index constituents corresponding to market m at time t in World Cup sub-sample w. I take the mean because the number of actively traded stocks in each market index can vary over time.

2.3.5. Adjustments

I make two adjustments to the trading activity and market condition variables presented in subsections 2.3.3 and 2.3.4. First, to ensure that each market variable has the same order of magnitude across the markets, I standardise each individual market-level series to have a mean of zero and standard deviation of one. This also controls for country fixed effects. Second, to control for seasonal trends in the data and isolate the abnormal trading activity resulting from the football matches, all data is adjusted using the methodology Gallant et al. (1992). Assuming a normal distribution, the seasonality adjustment process maintains the exact mean and variance of the data, after removing all variation that is explained by the seasonal variables. The seasonal variables consist of: month-of-the-year indicator variables; day-of-the-week indicator variables; and, five-minute time-of-the-day indicator variables.

2.3.6. Summary Statistics of the Market Condition Variables

Table 2.4 presents the summary statistics of the market condition variables: $AMIHUD_{m,t,w}$, $MAMIHUD_{m,t,w}$, $\sigma_{m,t,w}$ and $PD_{m,t,w}$. Due to the adjustments to the data described in the previous subsection, all market variables have a mean close to zero and a standard deviation close to one. Further, most variables are positively skewed for every World Cup sub-sample. The exception is the $AMIHUD_{m,t,w}$ measure which is negatively skews for the 2010 and 2014 World Cup sub-samples. Interestingly, the $MAMIHUD_{m,t,w}$ measure displays more positive skewness

Table 2.4
Summary Statistics of the Market Condition Variables.

This table presents the summary statistics of the market condition variations for each World Cup sub-sample w. The market condition variables include: $AMIHUD_{m,t,w}$, the mean Amihud (2002) measure for market m at time t and World Cup sub-sample w; $MAMIHUD_{m,t,w}$, the mean modified Amihud (2002) measure; $\sigma_{m,t,w}$, mean volatility; and, gross price discovery, denoted by $PD_{m,t,w}$.

Panel A: Raw Observations						
Year	$DEP_{m,t,w}$	Mean	Median	Standard Deviation	Skewness	Kurtosis
1998	$AMIHUD_{m,t,1998}$	-0.001	0.001	0.802	0.146	57.881
	$MAMIHUD_{m,t,1998}$	-0.001	-0.278	0.969	1.348	39.725
	$VOLATILITY_{m,t,1998}$	-0.004	-0.087	0.860	2.352	19.781
	$PD_{m,t,1998}$	0.001	-0.091	0.998	0.542	9.182
2002	$AMIHUD_{m,t,2002}$	0.001	0.008	0.832	0.814	69.034
	$MAMIHUD_{m,t,2002}$	0.025	-0.026	0.945	3.045	138.991
	$VOLATILITY_{m,t,2002}$	0.012	-0.173	0.909	3.453	41.681
	$PD_{m,t,2002}$	0.001	-0.074	0.992	0.475	6.228
2006	$AMIHUD_{m,t,2006}$	-0.001	0.001	0.935	0.057	16.881
	$MAMIHUD_{m,t,2006}$	0.021	-0.162	1.000	8.040	137.123
	$VOLATILITY_{m,t,2006}$	0.005	-0.099	0.932	1.680	9.425
	$PD_{m,t,2006}$	0.002	-0.082	0.997	0.690	7.816
2010	$AMIHUD_{m,t,2010}$	0.001	0.006	0.928	-0.220	32.325
	$MAMIHUD_{m,t,2010}$	0.023	-0.218	0.993	7.779	150.769
	$VOLATILITY_{m,t,2010}$	0.001	-0.114	0.823	2.796	27.109
	$PD_{m,t,2010}$	0.000	-0.076	0.998	0.672	6.052
2014	$AMIHUD_{m,t,2014}$	0.002	0.002	0.923	-0.501	25.105
	$MAMIHUD_{m,t,2014}$	0.022	-0.046	0.929	1.943	36.906
	$VOLATILITY_{m,t,2014}$	0.002	-0.166	0.912	1.933	15.537
	$PD_{m,t,2014}$	0.001	-0.059	1.000	0.305	7.506

than the $AMIHUD_{m,t,w}$ measure for every World Cup sub-sample and greater kurtosis in every World Cup sub-sample except the 1998 sub-sample. Table 2.4 demonstrates that every variable can be classified as leptokurtic, with $PD_{m,t,w}$ consistently having the least kurtosis across the World Cup sub-samples. To address concerns over the impact of outliers in the market condition variables, I also consider winsorising each variable at various thresholds and taking the logarithmic transformation of the $AMIHUD_{m,t,w}$, $MAMIHUD_{m,t,w}$ and $\sigma_{m,t,w}$ variables. The empirical results are robust to the logarithmic transformation and winsorisation. Hence, for simplicity, I present the results for the variables as defined in Sub-section 2.3.4.

2.4. Discretionary Trading on Match Days

In this section, I present evidence of abnormal discretionary trading on World Cup match days. Sub-section 2.4.1 presents graphical evidence. Sub-section 2.4.2 presents more rigorous statistical evidence in the form of a regression analysis.

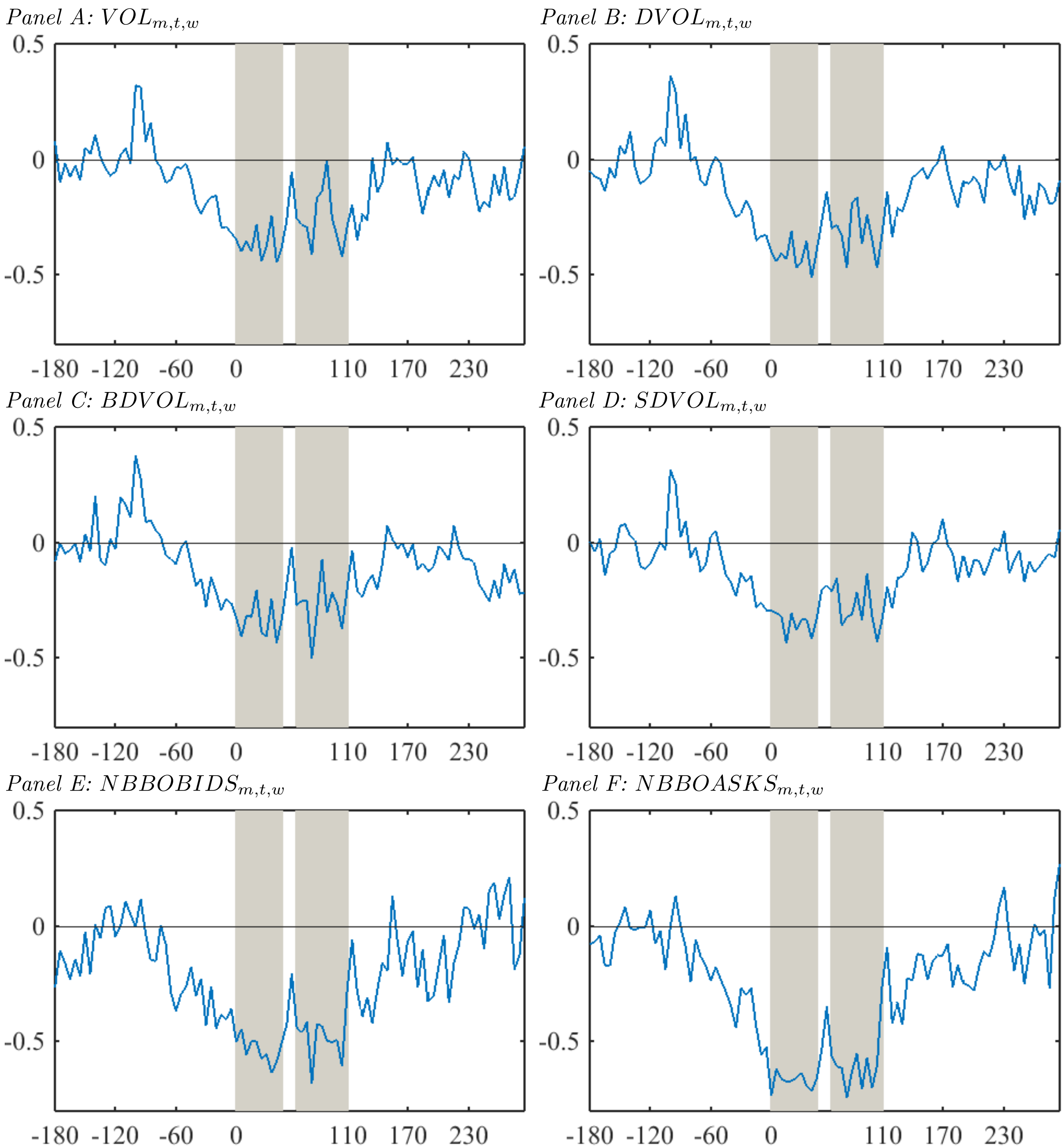


Fig. 2.1. Trading Activity On Match Days. This figure plots the mean standardised, seasonally detrended trading activity variables on match days. The market variables are detrended using the methodology of Gallant et al. (1992) for month-of-the-year, day-of-the-week and five-minute time-of-the-day effects. Volume for market m in World Cup sub-sample w at time t is denoted by $VOL_{m,t,w}$, dollar volume by $DVOL_{m,t,w}$, number of trades by $TRADES_{m,t,w}$, buyer-initiated dollar trading volume by $BDVOL_{m,t,w}$, seller-initiated dollar trading volume by $SDVOL_{m,t,w}$, number of bids at the national best bid price by $NBBOBIDS_{m,t,w}$ and number of asks at the national best ask price is denoted by $NBBOASKS_{m,t,w}$. The average first-half and second-half time periods are shaded in grey. The x-axis of each plot indicates the number of minutes from kick-off time.

2.4.1. Visual Analysis

Figure 2.1 plots the mean $VOL_{m,t,w}$, $DVOL_{m,t,w}$, $BDVOL_{m,t,w}$, $DVOL_{m,t,w}$, $NBBOBIDS_{m,t,w}$ and $NBBOASKS_{m,t,w}$ values on match day. Figure 2.1 is designed to allow for a simple visual interpretation of the data. The trading activity variables are normalised at the country level and are seasonally detrended. Therefore, if any of the variables of interest deviate from zero, it can be interpreted as abnormal market activity.

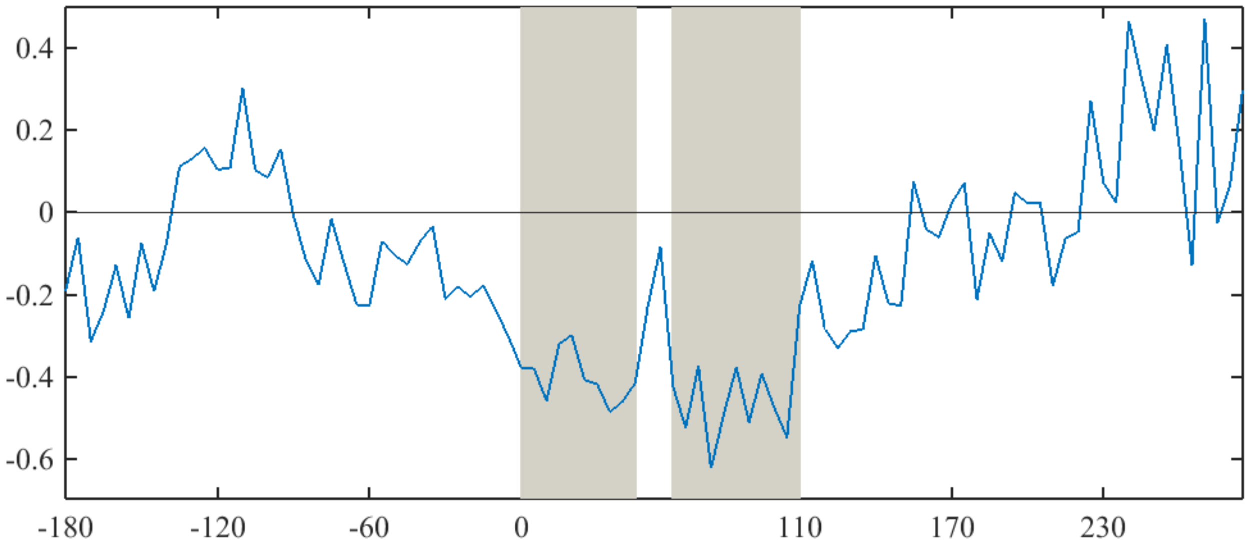
Figure 2.1 demonstrates that World Cup football matches have a very significant impact on the entire trading day. This is in contrast to Ehrmann and Jansen (2017) whom only consider the contemporaneous distraction effect of football matches. Before matches, there is a pronounced period of heightened trading between one to two hours before kick-off time. In Panel A, $VOL_{m,t,w}$ reaches a peak of 0.33 standard deviations above the mean, while in Panel B, $DVOL_{m,t,w}$ peaks at 0.41 standard deviations. Following this period, trading activity declines sharply towards kick-off time. Following kick-off, there is a sustained period of reduced trading activity during match time. This is consistent with the empirical results of Ehrmann and Jansen (2017). Match time is punctuated by a spike in trading at half-time. Following full-time, trading activity gradually increases towards normal levels but mostly remain below normal levels for the remainder of the trading day.

Panels C and D plot the $BDVOL_{m,t,w}$ and $SDVOL_{m,t,w}$ variables. Both $BDVOL_{m,t,w}$ and $SDVOL_{m,t,w}$ exhibit a similar trend over the trading day. This is consistent with the notion of liquidity trading and inconsistent with traders simply closing out their positions before kick-off.

Panels E and F plot the number of bids and asks at the NBBO, $NBBOBIDS_{m,t,w}$ and $NBBOASKS_{m,t,w}$ respectively. Interestingly, while $NBBOBIDS_{m,t,w}$ and $NBBOASKS_{m,t,w}$ decline substantially during match time, they do not increase as substantially as the other variables prior to match time. This suggests that liquidity providers do not substantially increase the depth of the limit order book during the pre-match period but instead *maintain* relatively normal levels of liquidity, despite the abnormal trading volume suggested by panels A, B, C and D.

Figure 2.1 provides very strong evidence that investors refrain from trading during important football matches. Further, the persistent spike in trading activity during the 15 minute half-time period suggests that some traders monitor football matches during match time and revert their attention back to the market during half-time. The positive abnormal trading levels from two to one hours before kick-off suggest a temporal substitution effect whereby market participants choose to avoid trading during match time and instead fulfil their trading requirements before the match starts.

Panel A: $BREVISIONS_{m,t,w}$



Panel B: $AREVISIONS_{m,t,w}$

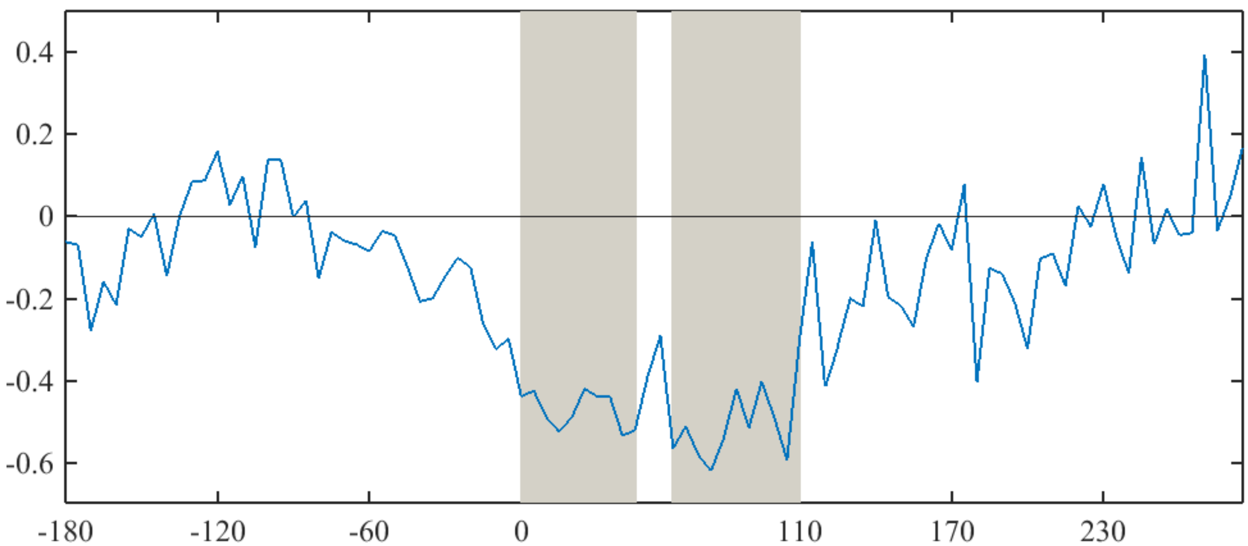


Fig. 2.2. Order Revisions and Cancellations at the NBBO on Match Days. This figure plots the mean standardised, seasonally detrended limit order revision variables on match days. The market variables are detrended using the methodology of Gallant et al. (1992) for month-of-the-year, day-of-the-week and five-minute time-of-the-day effects. The bid order revision proxy, $BREVISIONS_{m,t,w}$, counts the number of times the best bid price or volume of the best bid price for stocks in market m at time t in World Cup sub-sample w decrease without a trade taking place. The ask order revision proxy, $AREVISIONS_{m,t,w}$ counts the number of times the best ask price increases or volume of the best ask price decreases for stocks in market m at time t in World Cup sub-sample w decrease without a trade taking place. The average first-half and second-half time periods are shaded in grey. The x-axis of each plot indicates the number of minutes from kick-off time.

Following the period of heightened trading, trading activity begins to decline to abnormally low levels. This decline begins well before kick-off time. This can be explained by the distraction hypothesis of Ehrmann and Jansen (2017) and the discretionary trading hypothesis. Under the distraction hypothesis, the gradual decline in trading can be explained by a gradual increase in the number of traders being distracted. For example, in the lead up to kick-off, traders could be distracted by national anthems or pre-match analyses. Under the discretionary trading hypotheses, the gradual decline in trading activity can be explained by discretionary traders gradually leaving the market after satisfying their trading demands. Some market participants might also leave their trading desks well before kick-off time in order to commute to a viewing venue.

The reduced amount of trading following full-time can similarly be explained by distraction and discretionary trading. The reduced trading can be attributed to a loss in productivity due to the distraction of post-match analyses, celebrations or commiserations. The reduced level of trading can also be attributed to discretionary traders not returning to the market after full-time.

Figure 2.2 plots the mean $BREVISIONS_{m,t,w}$ and $AREVISIONS_{m,t,w}$ values on match day. The mean series follow the same trend as the $VOL_{m,t,w}$ and $DVOL_{m,t,w}$ variables presented in Figure 2.1. This is consistent with the model of Liu (2009). According to Liu (2009), those who place limit orders face two distinct risks: non-execution (NE) risk and free-option (FO) risk. NE risk is the risk of a limit order not resulting in a trade. FO risk is the risk of a limit order resulting in a trade at an unfavourable price. To mitigate these risks, traders monitor the market and cancel or revise their limit orders accordingly. Traders revise or cancel their orders more frequently when markets are more active because prices may move away from their limit price such that they lose price priority (NE risk) or their limit order could be picked off as new information enters the market (FO risk). Thus, Figure 2.2 is representative of the abnormal trading activity patterns presented in Figure 2.1 and the limit order submission risks during match days.

2.4.2. Regression Analysis

2.4.2.1 Daily Analysis

In this subsection, I test the null hypothesis that football matches do not have an impact on trading activity at the daily level. To do this, I estimate the following equation:

$$DEP_{m,t,w} = \alpha_0 + \beta_0 GD_{m,t,w} + \epsilon_{m,t,w} \tag{2.1}$$

where $DEP_{m,t,w}$ is the dependent variable for market m at time t and World Cup sub-sample w. The match day indicator variable, $GD_{m,t,w}$, takes the value of one if five-minute intra-day time period t coincides with a trading day in which country m is open for trading simultaneous to country m participating in a football match. Thus, if β_0 is significantly different from zero, we can conclude that there is an abnormal amount of trading activity that occurs on match days.

Table 2.5
The Daily Impact of Matches on Trading Activity

This table reports the estimation results for the following regression:

$$DEP_{m,t,w} = \alpha_0 + \beta_0 GD_{m,t,w} + \epsilon_{m,t,w} \tag{2.1}$$

where $DEP_{m,t,w}$ is the dependent variable for market m at time t and World Cup sub-sample w. Each dependent variable is formed by pooling the standardised, detrended market-level observations. The $GD_{m,t,w}$ indicator variable is a match day indicator variable that takes the value of one if five-minute intra-day time period t coincides with a trading day in which country m is open for trading simultaneous to country m participating in a football match. Volume for market m at time t and World Cup sub-sample w is denoted by $VOL_{m,t,w}$, dollar volume by $DVOL_{m,t,w}$, number of trades by $TRADES_{m,t,w}$, buyer-initiated dollar trading volume by $BDVOL_{m,t,w}$ and seller-initiated dollar trading volume by $SDVOL_{m,t,w}$. The estimated β_0 values are reported for the full sample of countries and the individual country sub-samples. The t-statistics are reported in italics. The standard errors are clustered at the country-year level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively.

$DEP_{m,t,w}$	$VOL_{m,t,w}$	$DVOL_{m,t,w}$	$TRADES_{m,t,w}$	$BDVOL_{m,t,w}$	$SDVOL_{m,t,w}$
Full Sample	-0.119	-0.131	-0.249	-0.076	-0.136
	-1.06	*-1.15*	*-1.35*	*-0.88*	*-1.22*
Argentina	-0.231***	-0.445***	-0.481***	-0.142**	-0.368***
	-3.85	*-11.90*	*-5.09*	*-2.32*	*-15.92*
Belgium	-0.038	-0.012	-0.054	0.052	-0.155
	-0.35	*-0.12*	*-0.29*	*0.73*	*-1.45*
Brazil	-0.396***	-0.439***	-0.600***	-0.305***	-0.402***
	-3.89	*-5.15*	*-4.13*	*-4.07*	*-4.72*
Chile	0.016	0.001	0.072	0.021	0.001
	0.35	*0.01*	*0.84*	*0.37*	*0.01*
Colombia	-0.014	-0.251***	-0.646***	-0.050	-0.101**
	-0.30	*-4.35*	*-5.27*	*-1.00*	*-2.22*
Denmark	0.384	0.456	0.341	0.353	0.379
	1.44	*1.49*	*0.89*	*1.44*	*1.16*
England	-0.028	-0.148	-0.283	-0.039*	-0.116
	-1.24	*-1.36*	*-1.42*	*-1.73*	*-1.06*
France	-0.166***	-0.096***	-0.291**	-0.144***	-0.103
	-26.61	*-2.80*	*-2.44*	*-4.15*	*-1.58*
Germany	-0.188***	-0.177***	-0.214***	-0.105**	-0.176***
	-3.52	*-3.70*	*-5.67*	*-2.20*	*-4.36*
Greece	-0.357***	-0.266***	-0.422***	-0.284***	-0.181**
	-5.02	*-2.77*	*-6.28*	*-5.00*	*-2.13*
Ireland	-0.304***	-0.291***	-0.388*	-0.197***	-0.185***
	-4.87	*-7.15*	*-1.70*	*-3.49*	*-6.71*
Italy	-0.211***	-0.234***	-0.418***	-0.161***	-0.081***
	-10.19	*-109.70*	*-9.67*	*-3.66*	*-348.01*
Mexico	-0.157	-0.166*	-0.306***	-0.154	-0.203**
	-1.46	*-1.69*	*-3.54*	*-1.44*	*-2.32*

Table 2.5
(continued)

$DEP_{m,t,w}$	$VOL_{m,t,w}$	$DVOL_{m,t,w}$	$TRADES_{m,t,w}$	$BDVOL_{m,t,w}$	$SDVOL_{m,t,w}$
Netherlands	0.052	0.031	-0.063	0.207***	-0.073
	0.91	*0.51*	*-0.34*	*3.48*	*-0.47*
Poland	-0.125*	-0.130	-0.175	-0.061	-0.029
	-1.69	*-1.04*	*-1.44*	*-0.46*	*-0.31*
Portugal	-0.030	-0.065***	-0.167***	-0.020	-0.038**
	-1.16	*-30.05*	*-15.76*	*-0.37*	*-2.51*
Russia	-0.038	-0.081**	-0.074	-0.040	-0.039
	-1.20	*-1.97*	*-1.00*	*-0.97*	*-1.13*
South Africa	-0.164	-0.201	-0.236	-0.135	-0.163
	-0.85	*-0.77*	*-1.14*	*-0.72*	*-0.78*
Spain	-0.031	-0.022	0.037	-0.122***	0.027
	-0.47	*-0.32*	*1.15*	*-3.25*	*0.40*
Switzerland	0.100**	0.087	0.120***	0.063	0.118
	2.16	*1.35*	*6.15*	*1.55*	*1.59*
Turkey	-0.069	-0.002	-0.067	-0.045	0.002
	-0.79	*-0.02*	*-0.76*	*-0.41*	*0.03*
United States	-0.058	-0.041	-0.043	-0.119	0.010
	-0.35	*-0.24*	*-0.16*	*-1.39*	*0.05*

I estimate Equation 2.1 with respect to the full sample by pooling the country-level dependent variables together and clustering errors at the country-year level. I also estimate Equation 2.1 with respect to the individual country sub-samples.

Table 2.5 presents the estimation results of Equation 2.1 for the $VOL_{m,t,w}$, $DVOL_{m,t,w}$, $TRADES_{m,t,w}$, $BDVOL_{m,t,w}$ and $SDVOL_{m,t,w}$ dependent variables. All significant estimated β_0 coefficients presented in Table 2.5 are negative except for the Swiss sub-sample and the β_0 coefficient corresponding to the Dutch $BDVOL_{m,t,w}$ dependent variable. The β_0 coefficients are negative and significant for all dependent variables for Argentina, Brazil, France, Germany, Greece, Ireland and Italy.

For the pooled sample, the estimated β_0 coefficients are not statistically different from zero. The results indicate that we cannot categorically conclude that World Cup football matches that occur during trading hours cause *daily* trading activity to decline. The following sub-section demonstrates that the impact of World Cup football matches is more complex than a uniform decline in trading activity across match days.

2.4.2.2 Intra-day Analysis

In this subsection, I test the null hypothesis that football matches do not influence intra-day trading activity levels. To do this, I estimate the following equation:

$$\begin{aligned} DEP_{m,t,w} = & \alpha_0 + \beta_0 GD_{m,t,w} + \sum_{\tau=1}^{6} PRE_{m,t,w}^{K-\tau}\beta_\tau + \beta_7 D_{m,t,w}^F + \beta_8 D_{m,t,w}^H + \beta_9 D_{m,t,w}^S \\ & + \beta_{10} D_{m,t,w}^E + \sum_{\tau=1}^{6} POST_{m,t,w}^{F+\tau}\beta_{10+\tau} + \epsilon_{m,t,w} \end{aligned} \tag{2.2}$$

where $DEP_{m,t,w}$ is the dependent variable. The $D_{m,t,w}^F$ indicator variable takes the value of one if country m is playing in the first-half of a football match at time t in World Cup sub-sample w. The $D_{m,t,w}^S$ indicator variable is the analogous variable for the second-half of a football match. The $D_{m,t,w}^H$ indicator variable takes the value of one if it is half-time at time t for a match involving country m in World Cup sub-sample w. The $D_{m,t,w}^E$ indicator variable takes the value of one if it is extra-time at time t for a match involving country m in World Cup sub-sample w. The remaining indicator variables capture abnormal trading activity outside of match time on match days. The $PRE_{m,t,w}^{K-\tau}$ indicator variables take the value of one if t is 30τ minutes before a first-half kick-off observation or between 30τ and $30(\tau-1)$ minutes before a first-half kick-off observation, given that country m is participating in the match in World Cup sub-sample w. The $POST_{m,t,w}^{F+\tau}$ indicator variables take the value of one if t is 30τ minutes after a full-time observation or between 30τ and $30(\tau+1)$ minutes after a full-time observation, given that country m is participating in the match in World Cup sub-sample w.[14] Equation 2.2 is simultaneously estimated for all markets using errors clustered at the country-year level.

Equation 2.2 represents a significant deviation from the regression approach of Ehrmann and Jansen (2017). Ehrmann and Jansen (2017) only consider the contemporaneous impact of football matches on markets and the impact on trading within the immediate vicinity of match time. Instead, Equation 2.2 allows for a comprehensive analysis of trading levels up to three hours either side of match time. Equation 2.2 is inspired by the regression model of Foster and Viswanathan (1993). Foster and Viswanathan (1993) test for intra-day trading patterns across the entire trading day by constructing hourly time-of-the-day indicator variables.

Table 2.6 Panel A presents the estimation results of Equation 2.2 for the $VOL_{m,t,w}$, $DVOL_{m,t,w}$, $TRADES_{m,t,w}$, $BDVOL_{m,t,w}$ and $SDVOL_{m,t,w}$ dependent variables. Starting with the match time variables, $D_{m,t,w}^F$, $D_{m,t,w}^S$ and $D_{m,t,w}^E$, we can see that football matches have a large negative impact on trading activity during match time. For every trading activity variable in Table 2.6 Panel A, the estimated β_7, β_9 and β_{10} values are negative and significantly different from zero with a 95%

[14] A previous version of this paper included indicator variables for goals scored and goals conceded for country m at time t in World Cup sub-sample w. These indicator variables are excluded from this version as they proved to have a largely insignificant marginal impact on trading activity.

Table 2.6
Trading Activity On Match Days for All World Cups

This table reports the estimation results for the following regression:

$$\begin{aligned}DEP_{m,t,w} =& \alpha_0 + \beta_0 GD_{m,t,w} + \sum_{\tau=1}^{6} PRE^{K-\tau}_{m,t,w}\beta_\tau + \beta_7 D^F_{m,t,w} + \beta_8 D^H_{m,t,w} + \beta_9 D^S_{m,t,w} \\ &+ \beta_{10} D^E_{m,t,w} + \sum_{\tau=1}^{6} POST^{F+\tau}_{m,t,w}\beta_{10+\tau} + \epsilon_{m,t,w} \end{aligned} \tag{2.2}$$

where $DEP_{m,t,w}$ is the dependent variable. Each dependent variable is formed by pooling the standardised, detrended market-level observations across World Cup sub-samples. The match day indicator variable, $GD_{m,t,w}$, takes the value of one if five-minute intra-day time period t coincides with a trading day in which country m is open for trading simultaneous to country m participating in a football match. The $D^F_{m,t,w}$ indicator variable takes the value of one if country m is playing in the first-half of a football match at time t in World Cup sub-sample w. The $D^S_{m,t,w}$ indicator variable is the analogous variable for the second-half of a football match. The $D^H_{m,t,w}$ indicator variable takes the value of one if it is half-time at time t for a match involving country m in World Cup sub-sample w. The $D^E_{m,t,w}$ indicator variable takes the value of one if it is extra-time at time t for a match involving country m in World Cup sub-sample w. The remaining indicator variables capture abnormal trading activity outside of match time on match days. The $PRE^{K-\tau}_{m,t,w}$ indicator variables take the value of one if t is 30τ minutes before a first-half kick-off observation or between 30τ and $30(\tau-1)$ minutes before a first-half kick-off observation, given that country m is participating in the match in World Cup sub-sample w. The $POST^{F+\tau}_{m,t,w}$ indicator variables take the value of one if t is 30τ minutes after a full-time observation or between 30τ and $30(\tau+1)$ minutes after a full-time observation, given that country m is participating in the match in World Cup sub-sample w. Volume for market m at time t during World Cup sub-sample w is denoted by $VOL_{m,t,w}$, dollar volume by $DVOL_{m,t,w}$, number of trades by $TRADES_{m,t,w}$, buyer-initiated dollar trading volume by $BDVOL_{m,t,w}$, seller-initiated dollar trading volume by $SDVOL_{m,t,w}$, number of bids at the national best bid price by $NBBOBIDS_{m,t,w}$, number of asks at the national best ask price by $NBBOASKS_{m,t,w}$, total number of bids by $BIDS_{m,t,w}$ and total number of asks is denoted by $ASKS_{m,t,w}$. The t-statistics are reported in italics. The standard errors are clustered at the country-year level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively. R^2_W is a weighted coefficient of determination that only utilises observations for which at least one non-constant independent variable is nonzero.

Panel A: Trades

$DEP_{m,t,w}$	$VOL_{m,t,w}$	$DVOL_{m,t,w}$	$TRADES_{m,t,w}$	$BDVOL_{m,t,w}$	$SDVOL_{m,t,w}$
$GD_{m,t,w}$	-0.079**	-0.100***	-0.113**	-0.072**	-0.081**
	-2.34	*-2.86*	*-2.10*	*-2.39*	*-2.27*
$PRE^{K-6}_{m,t,w}$	0.031	0.011	0.105	0.027	0.024
	0.46	*0.18*	*1.37*	*0.41*	*0.39*
$PRE^{K-5}_{m,t,w}$	0.078	0.082	0.127**	0.063	0.085**
	1.61	*1.56*	*2.06*	*0.89*	*2.31*
$PRE^{K-4}_{m,t,w}$	0.184**	0.234**	0.213*	0.251**	0.149*
	2.07	*2.42*	*1.80*	*2.25*	*1.94*
$PRE^{K-3}_{m,t,w}$	0.085	0.115	0.054	0.099	0.060
	1.17	*1.49*	*0.64*	*1.43*	*0.81*
$PRE^{K-2}_{m,t,w}$	-0.025	-0.013	-0.095	-0.022	-0.008
	-0.42	*-0.23*	*-0.97*	*-0.36*	*-0.17*

Table 2.6
(continued)

$DEP_{m,t,w}$	$VOL_{m,t,w}$	$DVOL_{m,t,w}$	$TRADES_{m,t,w}$	$BDVOL_{m,t,w}$	$SDVOL_{m,t,w}$
$PRE^{K-1}_{m,t,w}$	-0.151**	-0.163**	-0.316***	-0.152**	-0.130**
	-2.19	*-2.32*	*-2.95*	*-2.48*	*-2.06*
$D^{F}_{m,t,w}$	-0.280***	-0.308***	-0.593***	-0.259***	-0.255***
	-4.75	*-4.28*	*-6.24*	*-3.82*	*-4.28*
$D^{H}_{m,t,w}$	-0.101	-0.119	-0.258**	-0.068	-0.109*
	-1.37	*-1.64*	*-2.34*	*-0.77*	*-1.71*
$D^{S}_{m,t,w}$	-0.168**	-0.197***	-0.497***	-0.196***	-0.188***
	-2.39	*-2.96*	*-5.77*	*-2.96*	*-3.26*
$D^{E}_{m,t,w}$	-0.739***	-0.919***	-1.494***	-0.846***	-0.618***
	-19.51	*-11.30*	*-16.85*	*-28.41*	*-10.73*
$POST^{F+1}_{m,t,w}$	-0.117*	-0.095	-0.206**	-0.065	-0.089*
	-1.93	*-1.51*	*-2.30*	*-1.28*	*-1.67*
$POST^{F+2}_{m,t,w}$	0.060	0.047	0.110	0.040	0.046
	0.77	*0.86*	*1.32*	*0.81*	*0.80*
$POST^{F+3}_{m,t,w}$	-0.023	0.033	-0.014	0.017	0.019
	-0.31	*0.44*	*-0.12*	*0.29*	*0.25*
$POST^{F+4}_{m,t,w}$	0.058	0.081	0.140	0.078	0.068
	0.56	*0.73*	*0.89*	*0.79*	*0.81*
$POST^{F+5}_{m,t,w}$	0.076	0.053	0.093	-0.056	0.057
	0.57	*0.60*	*0.74*	*-0.79*	*0.67*
$POST^{F+6}_{m,t,w}$	-0.115*	-0.142*	-0.137	-0.154**	-0.050
	-1.66	*-1.65*	*-1.07*	*-2.03*	*-0.58*
Observations	291821	291821	291920	291824	291815
R^2_W (%)	2.42	3.24	6.79	2.77	2.36

Panel B: Bids and Asks

$DEP_{m,t,w}$	$NBBOBIDS_{m,t,w}$	$NBBOASKS_{m,t,w}$	$BIDS_{m,t,w}$	$ASKS_{m,t,w}$
$GD_{m,t,w}$	-0.147***	-0.141***	0.014	-0.037
	-2.87	*-2.60*	*0.31*	*-0.88*
$PRE^{K-6}_{m,t,w}$	-0.050	0.025	0.158*	0.056
	-0.77	*0.33*	*1.65*	*0.79*
$PRE^{K-5}_{m,t,w}$	0.113	0.135**	0.073	0.104
	1.63	*2.39*	*0.86*	*1.25*
$PRE^{K-4}_{m,t,w}$	0.172**	0.122	0.068	0.100
	2.08	*1.55*	*0.74*	*1.09*
$PRE^{K-3}_{m,t,w}$	0.022	0.016	0.011	0.091
	0.28	*0.17*	*0.14*	*1.12*
$PRE^{K-2}_{m,t,w}$	-0.137	-0.160	-0.021	-0.029
	-1.14	*-1.52*	*-0.34*	*-0.41*
$PRE^{K-1}_{m,t,w}$	-0.227*	-0.263**	-0.109	-0.061
	-1.75	*-2.26*	*-1.50*	*-0.86*
$D^{F}_{m,t,w}$	-0.396***	-0.537***	-0.303***	-0.246***
	-3.21	*-4.82*	*-4.72*	*-3.94*
$D^{H}_{m,t,w}$	-0.203	-0.339***	-0.182**	-0.158**
	-1.56	*-3.38*	*-2.11*	*-2.35*

Table 2.6
(continued)

$DEP_{m,t,w}$	$NBBOBIDS_{m,t,w}$	$NBBOASKS_{m,t,w}$	$BIDS_{m,t,w}$	$ASKS_{m,t,w}$
$D^{S}_{m,t,w}$	-0.354***	-0.496***	-0.223***	-0.182**
	-3.73	*-5.50*	*-2.70*	*-2.45*
$D^{E}_{m,t,w}$	-1.423***	-1.560***	-0.620***	-0.573***
	-27.94	*-12.36*	*-9.02*	*-8.48*
$POST^{F+1}_{m,t,w}$	-0.117	-0.113	-0.050	-0.011
	-1.33	*-1.17*	*-0.58*	*-0.12*
$POST^{F+2}_{m,t,w}$	0.038	-0.029	0.071	0.153
	0.48	*-0.37*	*0.48*	*1.10*
$POST^{F+3}_{m,t,w}$	-0.025	-0.051	-0.149**	0.016
	-0.20	*-0.40*	*-1.98*	*0.17*
$POST^{F+4}_{m,t,w}$	0.063	0.104	-0.080	-0.079
	0.40	*0.65*	*-0.69*	*-1.06*
$POST^{F+5}_{m,t,w}$	0.222	0.117	-0.182*	-0.152
	1.11	*0.93*	*-1.86*	*-1.44*
$POST^{F+6}_{m,t,w}$	0.095	0.072	-0.034	-0.006
	0.58	*0.60*	*-0.46*	*-0.09*
Observations	291922	291928	268694	263144
R^2_W (%)	3.40	5.56	1.90	1.68

level of confidence.[15] The magnitudes of the estimated β_7 and β_9 coefficients are very significant. The β_7 estimated coefficients in Table 2.6 Panel A range from -0.255 for $SDVOL_{m,t,w}$ to -0.593 for $TRADES_{m,t,w}$. This means that on average, $SDVOL_{m,t,w}$ decreases by 25.5% of the standard deviation and $TRADES_{m,t,w}$ decreases by 59.3% of the standard deviation during the first half of a football match. The extra-time indicator variable, $D^{E}_{m,t,w}$, has the greatest impact on trading activity. This is unsurprising as extra-time can only occur for those matches within the "elimination-stage" of a World Cup. These matches are perceived to be of greater importance because the loser is immediately eliminated from the World Cup and because these matches occur towards the end of the tournament when only a few countries are yet to be eliminated. The estimated β_{10} coefficients associated with the $D^{E}_{m,t,w}$ variable range from -0.618 for $SDVOL_{m,t,w}$to -1.494 for $TRADES_{m,t,w}$. Thus, trades drop by 1.494 standard deviations during extra-time of a World Cup match.

Table 2.6 Panel A also shows that trading in the immediate vicinity of match time is significantly reduced. The estimated β_6 coefficients are negative and statistically different from zero for all trading activity variables. Further, the estimated β_{11} coefficients are negative and statistically significant from zero for the $VOL_{m,t,w}$, $TRADES_{m,t,w}$ and $SDVOL_{m,t,w}$ dependent variables. The β_6 and β_{11} estimation results are consistent with Ehrmann and Jansen (2017). Further, all significant coefficients relating to the $POST^{F+\tau}_{m,t,w}$ variables in Table 2.6 Panel A are negative. Thus, trading activity is generally reduced in the post-match period.

The pre-match period contrasts significantly to the post-match period. In Table 2.6 Panel A,

[15]Ehrmann and Jansen (2017) only consider the number of trades and trading volume in their study. Thus, the estimated β_7 and β_9 coefficients for $VOL_{m,t,w}$ and $TRADES_{m,t,w}$ confirm the main result of Ehrmann and Jansen (2017).

all estimated β_τ coefficients for $\tau \in \{3, 4, 5, 6\}$ are positive while all estimated β_τ coefficients for $\tau \in \{1, 2\}$ are negative. Thus, in the third and second last hour before kick-off, trading activity is increased while in the last hour before kick-off trading activity is reduced.

Table 2.6 Panel A also confirms the statistical significance of the positive abnormal level of discretionary trading in the second last hour before kick-off. The highest level of trading occurs between 120 to 90 minutes before kick-off. For all trading activity variables in Table 2.6 Panel A, the estimated β_4 coefficients have the greatest magnitude. From 120 to 90 minutes until kick-off, the number of trades is 21.3% of a standard deviation higher than normal; while, dollar trading volume is 23.4% of a standard deviation higher than normal. Thus, we can reject the null hypothesis that football matches do not impact trading activity during match time and outside of match time.

Panel B of Table 2.6 gives the Equation 2.2 estimation results for the bid and ask variables: $NBBOBIDS_{m,t,w}$, $NBBOASKS_{m,t,w}$, $BIDS_{m,t,w}$ and $ASKS_{m,t,w}$. With respect to the match-time variables, $D^F_{m,t,w}$, $D^S_{m,t,w}$ and $D^E_{m,t,w}$, all estimated coefficients are negative and significant at the 95% level of confidence. Moreover, in line with the trading activity variables in Panel A, the extra-time indicator variable, $D^E_{m,t,w}$, demands the most negative coefficient. In the pre-match period, there is also some evidence of a greater number of bids and asks in the limit order book. In particular, all estimated β_τ coefficients for $\tau \in \{3, 4, 5\}$ are positive. Nonetheless, the increase in bids and asks is during the pre-match period of the trading day relatively insignificant. Finally, all estimated coefficients for the period after full-time are negative. This suggests that there are fewer limit orders submissions following World Cup matches.

It should also be noted that the explanatory power of Equation 2.2 is quite low. This is because the match day observations only constitute a small proportion of the total sample. Since the explanatory power of Equation 2.2 is not of key interest, R-squared statistics are not shown. Instead, Table 2.6 gives a weighted R^2 coefficient that only utilises observations for which at least one non-constant independent variable is nonzero.

2.5. Market Conditions on Match Days

This section examines market conditions during World Cup match days. Sub-section 2.5.1 describes the Admati and Pfleiderer (1988) predictions for market conditions with respect to discretionary trading. Sub-section 2.5.2 plots the market condition variables on match days. Sub-section 2.5.3 formally tests for abnormal market conditions on match days in a regression setting.

2.5.1. The Admati and Pfleiderer (1988) Testable Hypotheses

In their influential paper, Admati and Pfleiderer (1988) provide a theoretical explanation for observed trading patterns. Specifically, Admati and Pfleiderer (1988) seek to explain why trading and return volatility are amplified at particular times during the trading day and why trading volume is correlated to return variability. Admati and Pfleiderer (1988) demonstrate that these observed trading patterns can be explained by two groups of optimising agents: discretionary liquidity traders (DLTs) and endogenously informed traders. In their baseline model, Admati and Pfleiderer (1988) extend on the Kyle (1985) model by describing two types of liquidity traders: non-discretionary liquidity traders (NDLTs) and DLTs. NDLTs are uninformed and are randomly assigned a net demand that must be satisfied in each period. On the other hand, DLTs decide when to trade. In the Admati and Pfleiderer (1988) model, there are T periods. In period T', each DLT's net demand is determined. This net demand needs to be satisfied before T'', where $T' < T'' < T$. In the baseline model, each DLT can only trade once. Admati and Pfleiderer (1988) demonstrate that in this setting, there will always exist equilibria in which all DLTs trade in the same period. They trade in a single period, when market depth is at its greatest and transaction costs at their lowest. The influx of discretionary liquidity trading also encourages informed traders to increase their demand in this period. The increase in informed trading in this period has the effect of increasing price variance, as price moves to reflect the new information contained in the informed order flow. This gives the first testable hypothesis of the Admati and Pfleiderer (1988) model:

H1 When there is increased discretionary trading:

a. Price impact costs are *reduced*;
b. Volatility is *increased*; and,
c. Price discovery is *increased*.

Conversely, during periods in which informed traders only trade with NDLTs, adverse selection costs are higher and price variance is lower. This gives the second testable hypothesis of the Admati and Pfleiderer (1988) model:

H2 When there is decreased discretionary trading:

a. Price impact costs are *increased*;
b. Volatility is *reduced*; and,
c. Price discovery is *reduced*.

Under Hypothesis H1, the period of increased discretionary trading in the second last hour before the kick-off of a football match should be accompanied by *reduced* price impact costs, *increased* volatility and *increased* price discovery. Under Hypothesis H2, the periods in which a football match is being played and there is reduced discretionary trading should be accompanied by *increased* price impact costs, *reduced* volatility and *reduced* price discovery.

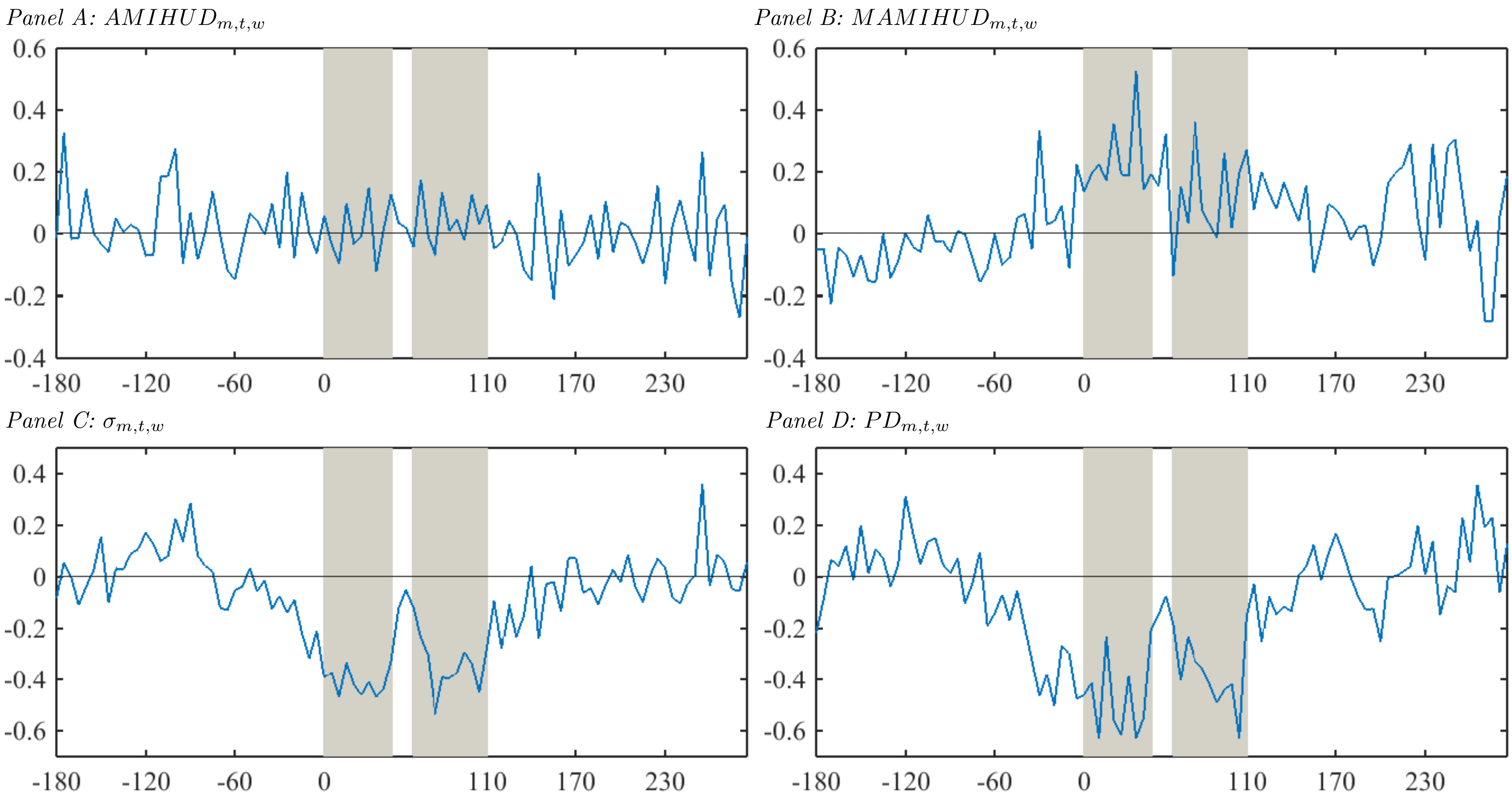


Fig. 2.3. Market Conditions On Match Days. This figure plots the mean standardised, seasonally detrended market condition variables on match days. The market variables are detrended using the methodology of Gallant et al. (1992) for month-of-the-year, day-of-the-week and five-minute time-of-the-day effects. The mean Amihud (2002) measure for market m at time t and World Cup sub-sample w is denoted by $AMIHUD_{m,t,w}$, the mean modified Amihud (2002) measure by $MAMIHUD_{m,t,w}$, mean volatility by $\sigma_{m,t,w}$ and gross price discovery by $PD_{m,t,w}$. The average first-half and second-half time periods are shaded in grey. The x-axis of each plot indicates the number of minutes from kick-off time.

2.5.2. *Visual Analysis*

Figure 2.3 plots the mean market condition variables on match days. Panels A and B plot the price impact measures, $AMIHUD_{m,t,w}$ and $MAMIHUD_{m,t,w}$. Panel A demonstrates that the $AMIHUD_{m,t,w}$ measure does not display any distinct intra-day trend on match days. In contrast, Panel B shows that the $MAMIHUD_{m,t,w}$ measure displays a compelling intra-day trend on match days. In line with Hypothesis H1.a, the $MAMIHUD_{m,t,w}$ measure mostly remains below zero during the pre-match period. Further, in line with H2.a, the $MAMIHUD_{m,t,w}$ measure is mostly increased during match time. The mean $MAMIHUD_{m,t,w}$ value reaches a peak of 0.53 standard deviations above normal levels. Following match time, the $MAMIHUD_{m,t,w}$ measure approaches normal levels.

The $\sigma_{m,t,w}$ and $PD_{m,t,w}$ market variables display a different intra-day pattern to the $MAMIHUD_{m,t,w}$ measure. Panel C plots the mean $\sigma_{m,t,w}$ series around match time. Panel C shows that volatility displays a U-shaped pattern around match time with a distinct spike during half-time. Consistent with the Admati and Pfleiderer (1988) predictions, there is increased volatility during the concentrated trading period, between two to one hours before kick-off, and there is decreased volatility during matches. Hence, there is evidence in favour of both Hypothesis H1.b and Hypothesis H2.b. Figure 2.3 Panel D plots the mean $PD_{m,t,w}$ values on match days. Panel D shows that there is increased gross price discovery in the pre-match period of increased discretionary trading and a reduction in gross price discovery during match time. Further, Panel D shows a pronounced spike in price discovery during the half-time period. Hence, Figure 2.3 also provides evidence in favour of hypotheses H1.c and H2.c. Thus, the Admati and Pfleiderer (1988) predictions describe the volatility and gross price discovery conditions that occur on match days.

2.5.3. *Regression Analysis*

This section employs Equation 2.2 to test for abnormal market conditions during days in which a World Cup match occurs during trading hours. Table 2.7 presents the estimation results of Equation 2.2 with respect to the dependent variables: $AMIHUD_{m,t,w}$, $MAMIHUD_{m,t,w}$, $\sigma_{m,t,w}$ and $PD_{m,t,w}$.

The first dependent variables presented in Table 2.7 are the price impact measures: the $AMIHUD_{m,t,w}$ and $MAMIHUD_{m,t,w}$ measures. Under Hypotheses H1.a, price impact should be lower in the pre-match period of greater discretionary trading. Consistent with Figure 2.3, the estimation results only find significant evidence of reduced price impact costs during the pre-match period for the $MAMIHUD_{m,t,w}$ measure. For the $MAMIHUD_{m,t,w}$ measure, all significant β_τ estimated coefficients where $\tau \in \{2, 3, 4, 5, 6\}$ are negative. Thus, in line with Hypothesis H1.a, there is some evidence that price impact costs are reduced in the pre-matchperiod of increased discretionary trading. The pre-match estimated β_τ coefficients are not significantly different from zero for the $AMIHUD_{m,t,w}$ measure.

Table 2.7
Market Conditions On Match Days for All World Cups

This table reports the estimation results for the following regression:

$$DEP_{m,t,w} = \alpha_0 + \beta_0 GD_{m,t,w} + \sum_{\tau=1}^{6} PRE^{K-\tau}_{m,t,w}\beta_\tau + \beta_7 D^F_{m,t,w} + \beta_8 D^H_{m,t,w} + \beta_9 D^S_{m,t,w} + \beta_{10} D^E_{m,t,w} + \sum_{\tau=1}^{6} POST^{F+\tau}_{m,t,w}\beta_{10+\tau} + \epsilon_{m,t,w} \quad (2.2)$$

where $DEP_{m,t,w}$ is the dependent variable. Each dependent variable is formed by pooling the standardised, detrended market-level observations across World Cup sub-samples. The match day indicator variable, $GD_{m,t,w}$, takes the value of one if five-minute intra-day time period t coincides with a trading day in which country m is open for trading simultaneous to country m participating in a football match. The $D^F_{m,t,w}$ indicator variable takes the value of one if country m is playing in the first-half of a football match at time t in World Cup sub-sample w. The $D^S_{m,t,w}$ indicator variable is the analogous variable for the second-half of a football match. The $D^H_{m,t,w}$ indicator variable takes the value of one if it is half-time at time t for a match involving country m in World Cup sub-sample w. The $D^E_{m,t,w}$ indicator variable takes the value of one if it is extra-time at time t for a match involving country m in World Cup sub-sample w. The remaining indicator variables capture abnormal trading activity outside of match time on match days. The $PRE^{K-\tau}_{m,t,w}$ indicator variables take the value of one if t is 30τ minutes before a first-half kick-off observation or between 30τ and $30(\tau-1)$ minutes before a first-half kick-off observation, given that country m is participating in the match in World Cup sub-sample w. The $POST^{F+\tau}_{m,t,w}$ indicator variables take the value of one if t is 30τ minutes after a full-time observation or between 30τ and $30(\tau+1)$ minutes after a full-time observation, given that country m is participating in the match in World Cup sub-sample w. The mean Amihud (2002) measure for market m at time t and World Cup sub-sample w is denoted by $AMIHUD_{m,t,w}$, the mean modified Amihud (2002) measure by $MAMIHUD_{m,t,w}$, the mean volatility by $\sigma_{m,t,w}$ and the mean gross price discovery by $PD_{m,t,w}$. The t-statistics are reported in italics. The standard errors are clustered at the country level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively. R^2_W is a weighted coefficient of determination that only utilises observations for which at least one non-constant independent variable is nonzero.

$DEP_{m,t,w}$	$AMIHUD_{m,t,w}$	$MAMIHUD_{m,t,w}$	$\sigma_{m,t,w}$	$PD_{m,t,w}$
$GD_{m,t,w}$	-0.004	0.020	-0.086*	-0.060
	-0.22	*0.43*	*-1.80*	*-0.90*
$PRE^{K-6}_{m,t,w}$	0.074	-0.134*	0.063	0.006
	1.00	*-1.88*	*0.97*	*0.05*
$PRE^{K-5}_{m,t,w}$	0.005	-0.142**	0.124*	0.117
	0.11	*-2.09*	*1.89*	*0.96*
$PRE^{K-4}_{m,t,w}$	0.072	-0.053	0.215***	0.185*
	1.62	*-1.01*	*2.94*	*1.94*
$PRE^{K-3}_{m,t,w}$	0.005	-0.106**	0.105	0.041
	0.13	*-2.48*	*1.61*	*0.49*
$PRE^{K-2}_{m,t,w}$	0.007	-0.058	0.037	-0.107
	0.25	*-1.17*	*0.44*	*-0.94*
$PRE^{K-1}_{m,t,w}$	0.025	0.056	-0.087	-0.347
	0.63	*0.81*	*-0.95*	*-1.49*

Table 2.7
(continued)

$DEP_{m,t,w}$	$AMIHUD_{m,t,w}$	$MAMIHUD_{m,t,w}$	$\sigma_{m,t,w}$	$PD_{m,t,w}$
$D^F_{m,t,w}$	0.017	0.187**	-0.327***	-0.417***
	0.55	*2.19*	*-3.90*	*-3.53*
$D^H_{m,t,w}$	0.007	0.072	-0.011	-0.079
	0.15	*0.48*	*-0.11*	*-0.73*
$D^S_{m,t,w}$	0.057**	0.091	-0.281***	-0.335***
	2.23	*1.19*	*-4.05*	*-3.23*
$D^E_{m,t,w}$	-0.226	1.118***	-0.906***	0.051
	-1.50	*4.05*	*-15.10*	*0.04*
$POST^{F+1}_{m,t,w}$	-0.025	0.057	-0.101	-0.033
	-0.70	*0.76*	*-1.22*	*-0.31*
$POST^{F+2,w}_{m,t}$	-0.014	-0.037	0.029	0.098
	-0.20	*-0.50*	*0.40*	*0.88*
$POST^{F+3}_{m,t,w}$	0.003	-0.029	0.030	0.038
	0.10	*-0.32*	*0.28*	*0.31*
$POST^{F+4}_{m,t,w}$	-0.008	0.132	0.148	0.227
	-0.26	*1.17*	*1.07*	*1.52*
$POST^{F+5}_{m,t,w}$	0.010	0.122	0.167	0.113
	0.29	*1.10*	*1.11*	*0.59*
$POST^{F+6}_{m,t,w}$	-0.026	-0.132	0.061	0.009
	-0.31	*-1.07*	*0.55*	*0.07*
Observations	243310	248395	291870	243426
R^2_W (%)	0.12	1.22	3.09	1.92

As predicted by Hypothesis H2.a, price impact costs increase during match time. For both price impact measures, all significant coefficients relating to match time are positive. The estimated β_7 and β_{10} coefficients are both significant at the 95% confidence level for the $MAMIHUD_{m,t,w}$ dependent variable. The β_{10} coefficient indicates that the $MAMIHUD_{m,t,w}$ measure is on average 1.118 standard deviations higher during extra-time periods of a football match. Thus, the regression results provide strong support for Hypothesis H2.a and the notion that discretionary traders resist trading during match time to avoid high price impact costs.

The last two columns of Table 2.7 present the $\sigma_{m,t,w}$ and $PD_{m,t,w}$ estimation results. Under Hypothesis H1, $\sigma_{m,t,w}$ and $PD_{m,t,w}$ should be increased during the period of concentrated trades. Table 2.7 shows that the increase in volatility before match-time presented in Figure 2.3 Panel C is significant, particularly from 150 to 90 minutes before kick-off time. This coincides with the abnormal trading levels observed in Table 2.6. From 120 to 90 minutes before kick-off time, volatility is increased by 21.5% of a standard deviation. Table 2.6 also demonstrates that the increase in price discovery before match time observed in Figure 2.3 Panel D is statistically significant at the 90% level of confidence for the period of heightened discretionary trading, 120 to 90 minutes before kick-off time.

During match time, volatility is reduced and there is less gross price discovery. All estimated β_7, β_9 and β_{10} coefficients presented in the last two columns of Table 2.7 are negative and statistically significant at the 99% level of confidence, except the estimated β_{10} coefficient for $PD_{m,t,w}$. Volatility is lowest during the extra-time period of football matches. The $\sigma_{m,t,w}$ variable is reduced by 0.91 standard deviations during extra-time Thus, the results for $\sigma_{m,t,w}$ and $PD_{m,t,w}$ are consistent with hypotheses H1 and H2, as well as the assertion that increased discretionary trading induces increased trading by informed market participants.

2.6. Robustness

This section extends the analysis to demonstrate the robustness of the previous findings. Sub-section 2.6.1 examines trading activity on match days for a sample of small market capitalisation stocks. Sub-section 2.6.2 distinguishes between elimination-stage and non-elimination-stage football matches.

2.6.1. Small Market Capitalisation Sample

The main analysis focused on the constituents of the most important national market indices of the sample countries. These stocks have large market capitalisations. This sub-section analyses the trading activity and market conditions of small market capitalisation or "small cap" stocks. In this sub-section, the stock sample used in the main analysis, will be referred to as the "large cap" sample.

I select a small cap stock sample in an analogous manner to the large cap sample. For each country and World Cup iteration, a representative sample of small cap stocks is sampled from the constituents of a national small cap index. Table 2.8 outlines the national small cap indices. There are occasions where an analogous small cap index could not be found. Table 2.8 indicates these instances by "N/A". For example, in 2010 Argentina and Portugal do not have small cap market indices and are excluded from the small cap analysis. Further, to reduce computation time, 100 constituents of the S&P 600 Small Cap index are randomly drawn to represent small stocks from the United States of America.

I construct trading activity and market condition variables for the small cap sample using the same methodology implemented in the main analysis. The superscript 's' is used to differentiate the small cap market variables from the large cap variables that appear in the main analysis.

Figure 2.4 plots the mean small cap trading activity variables on match days. Figure 2.4 is analogous to Figure 2.1 in the main analysis. During match time, the small cap sample behaves in a similar fashion to the large cap sample. During the first and second half periods, small cap trading activity is markedly reduced. This is consistent with the notion that many traders of both large cap and small cap stocks do not to trade during World Cup football matches. Furthermore, as in the large cap sample, there is a spike in small cap trading activity during the 15-minute half-time period.

Table 2.8
Countries and Small Cap Market Indices

This table gives the small cap market indices that are used to construct the small cap stock sample. There are a number of instances whereby a small cap market index is not available. These instances are indicated by "N/A".

Country	1998	2002	2006	2010	2014
Argentina	N/A		N/A	N/A	N/A
Belgium	N/A	N/A			
Brazil	N/A		N/A	SMLL	
Chile				IGPAS	IGPAS
Colombia					N/A
Denmark	N/A			OMXCSCPI	
England	N/A	FTSC		FTSC	
France	N/A	N/A		CACS	
Germany		SDAXI	SDAXI	SDAXI	
Greece				N/A	
Ireland		ISCI			
Italy			FTITSC	FTITSC	
Mexico	N/A		N/A	MXXSM	MXXSM
Netherlands	N/A			ASCX	
Poland		N/A	N/A		
Portugal		N/A	N/A	N/A	
Russia		N/A			
South Africa				JSMLC	
Spain		N/A	IBEXS	IBEXS	
Switzerland			SSCC	SSCC	
Turkey		N/A			
Ukraine			N/A		
United States	N/A		SML	SML	SML

Figure 2.4 demonstrates that pre-match small cap trading differs from large cap pre-match trading. Figure 2.4 does not feature a great amount of discretionary trading prior to kick-off. This could be representative of the characteristics of small cap stocks. Small cap stocks attract investors with longer investment horizons (Stoll and Whaley (1983)). Stoll and Whaley (1983) demonstrate that due to the high transaction costs of small cap stocks, positive net small cap returns are best achieved through long investment horizons. In their sample, the minimum investment horizon to achieve a positive abnormal net return on a portfolio of small cap stocks is four months. This means that for small cap stocks, the 110 minute window of decreased liquidity during match time is unlikely to prompt a pre-match increase in discretionary trading caused by small cap investors with immediacy demands. Thus, the discrepancy between the two samples can be explained by a lack of demand for immediacy for small cap stocks that attract more long-term investors.

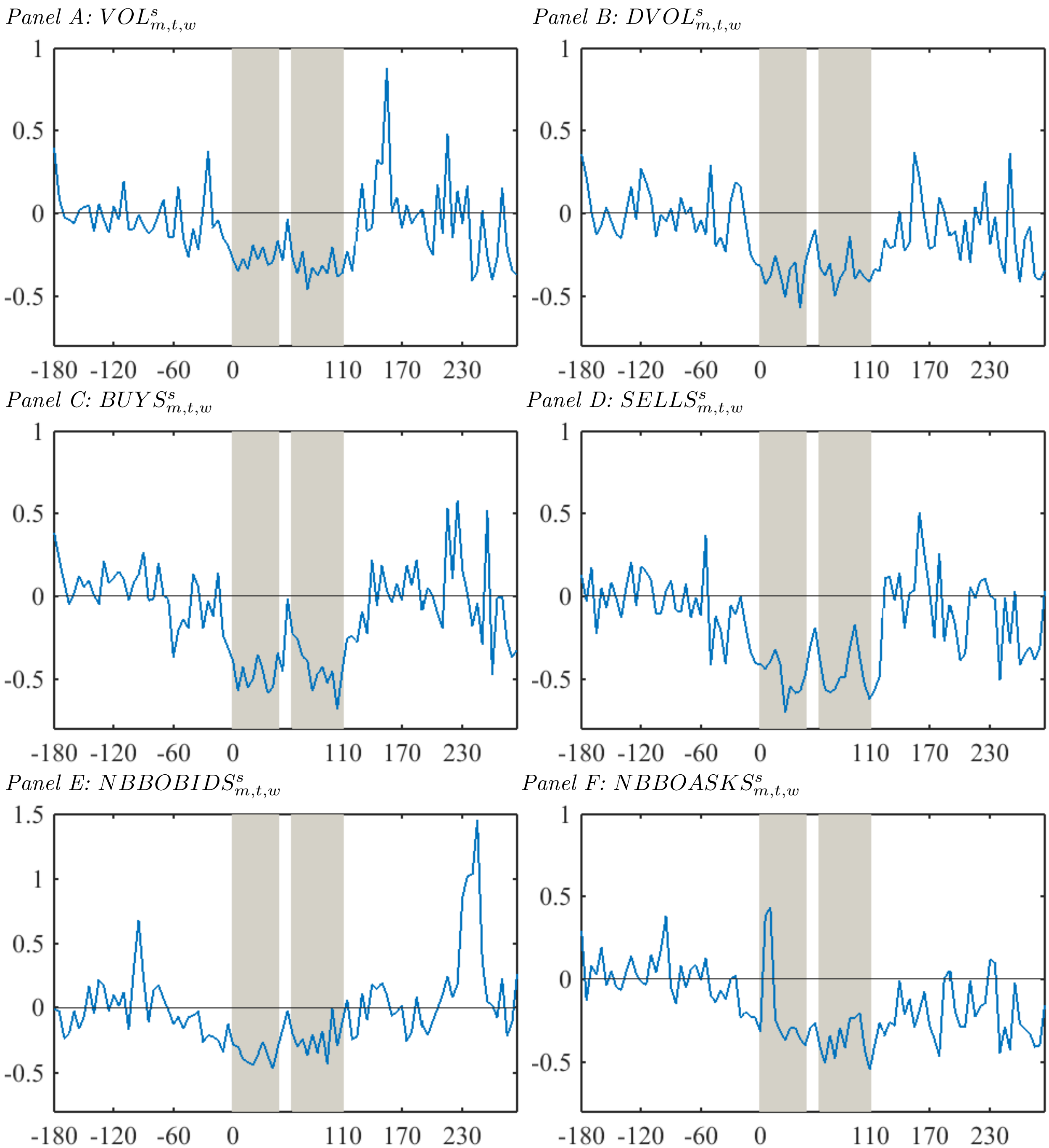


Fig. 2.4. Small Cap Trading Activity On Match Days. This figure plots the mean standardised, seasonally detrended small cap trading activity variables on match days. The market variables are detrended using the methodology of Gallant et al. (1992) for month-of-the-year, day-of-the-week and five-minute time-of-the-day effects. Small Cap volume for market m at time t is denoted by $VOL^s_{m,t,w}$, dollar volume by $DVOL^s_{m,t,w}$, number of trades by $TRADES^s_{m,t,w}$, buyer-initiated dollar trading volume by $BDVOL^s_{m,t,w}$, seller-initiated dollar trading volume by $SDVOL^s_{m,t,w}$, number of bids at the national best bid price by $NBBOBIDS_{m,t,w}$ and number of asks at the national best ask price is denoted by $NBBOASKS_{m,t,w}$. The average first-half and second-half time periods are shaded in grey. The x-axis of each plot indicates the number of minutes from kick-off time.

To determine whether World Cup football matches have a statistically significant effect on small cap trading activity, I estimate Equation 2.2 with respect to the small cap trading activity variables.[16] Table 2.9 Panel A presents the estimation results of Equation 2.2 for $VOL^s_{m,t,w}$, $DVOL^s_{m,t,w}$, $TRADES^s_{m,t,w}$, $BDVOL^s_{m,t,w}$ and $SDVOL^s_{m,t,w}$. Table 2.9 Panel A demonstrates that trading levels during match time are significantly lower than normal. For all trading variables, the estimated β_7 and β_9 coefficients are negative and statistically significant. The coefficients corresponding to the pre-match period confirm that the pre-match discretionary trading is less pronounced for the small cap sample. Nonetheless, all significant β_τ coefficients for $\tau \in \{4, 5, 6, 7\}$ are positive in Table 2.9 Panel A. Thus, Table 2.9 Panel A demonstrates that the small cap sample behaves in a similar fashion to the large cap sample with the exception of having a less significant amount of discretionary trading in the pre-match period between market open and kick-off time.[17]

Figure 2.5 plots the mean $AMIHUD^s_{m,t,w}$, $MAMIHUD^s_{m,t,w}$, $\sigma^s_{m,t,w}$ and $PD^s_{m,t,w}$ variables on match days. While the mean $AMIHUD^s_{m,t,w}$ and $MAMIHUD^s_{m,t,w}$ variables do not show any distinct intra-day pattern on match days, the $\sigma^s_{m,t,w}$ and $PD^s_{m,t,w}$ variables share a common trend with their large cap analogues. Prior to match time, $\sigma^s_{m,t,w}$ and $PD^s_{m,t,w}$ are generally increased relative to the rest of the trading day. During match time $\sigma^s_{m,t,w}$ decreases to a minimum of -0.48, while $PD^s_{m,t,w}$ decreases to -0.71. Thus, Figure 2.5 conforms to the H1.b, H1.c, H2.b and H2.c hypotheses

Table 2.9 Panel B presents the estimation results of Equation 2.2 with respect to the small cap market condition variables: $AMIHUD^s_{m,t,w}$, $MAMIHUD^s_{m,t,w}$, $\sigma^s_{m,t,w}$ and $PD^s_{m,t,w}$. In line with the relatively limited increase to trading in the pre-match period of match days for small cap stocks, Table 2.9 Panel B demonstrates that small cap market conditions prior to football matches are mostly insignificantly different from normal levels. During match time, the small cap market conditions follow a similar trend to the large cap market condition variables. Table 2.9 Panel B shows that all the estimated β_7 and β_9 values are of the sign that is predicted by the Admati and Pfleiderer (1988) model. The results for the $\sigma^s_{m,t,w}$ dependent variable indicate that volatility is reduced by 32.5% of a standard deviation during the first-half of a football match and 28.3% of a standard deviation during the second half of a football match. Further, the results for $PD^s_{m,t,w}$ indicate that price discover is reduced by 37.9% of a standard deviation for small cap stocks during the first-half of a football match. Despite this, the results during match time are weaker than those of the large cap sample. Thus, while the abnormal amount of discretionary trading that occurs in the pre-match period is less pronounced for the small cap sample, the small cap analysis nonetheless confirms the validity and robustness of the main results of interest.

[16] It should be noted that the countries with small cap data available did not feature in any extra-time observations that occurred during trading hours. Hence the $D^E_{m,t,w}$ variable is irrelevant for the small cap sample.

[17] For brevity, I do not present the Equation 2.2 estimation results for the $NBBOBIDS^s_{m,t,w}$, $NBBOASKS^s_{m,t,w}$, $BIDS^s_{m,t,w}$ nor $ASKS^s_{m,t,w}$ variables; however, they are available on request.

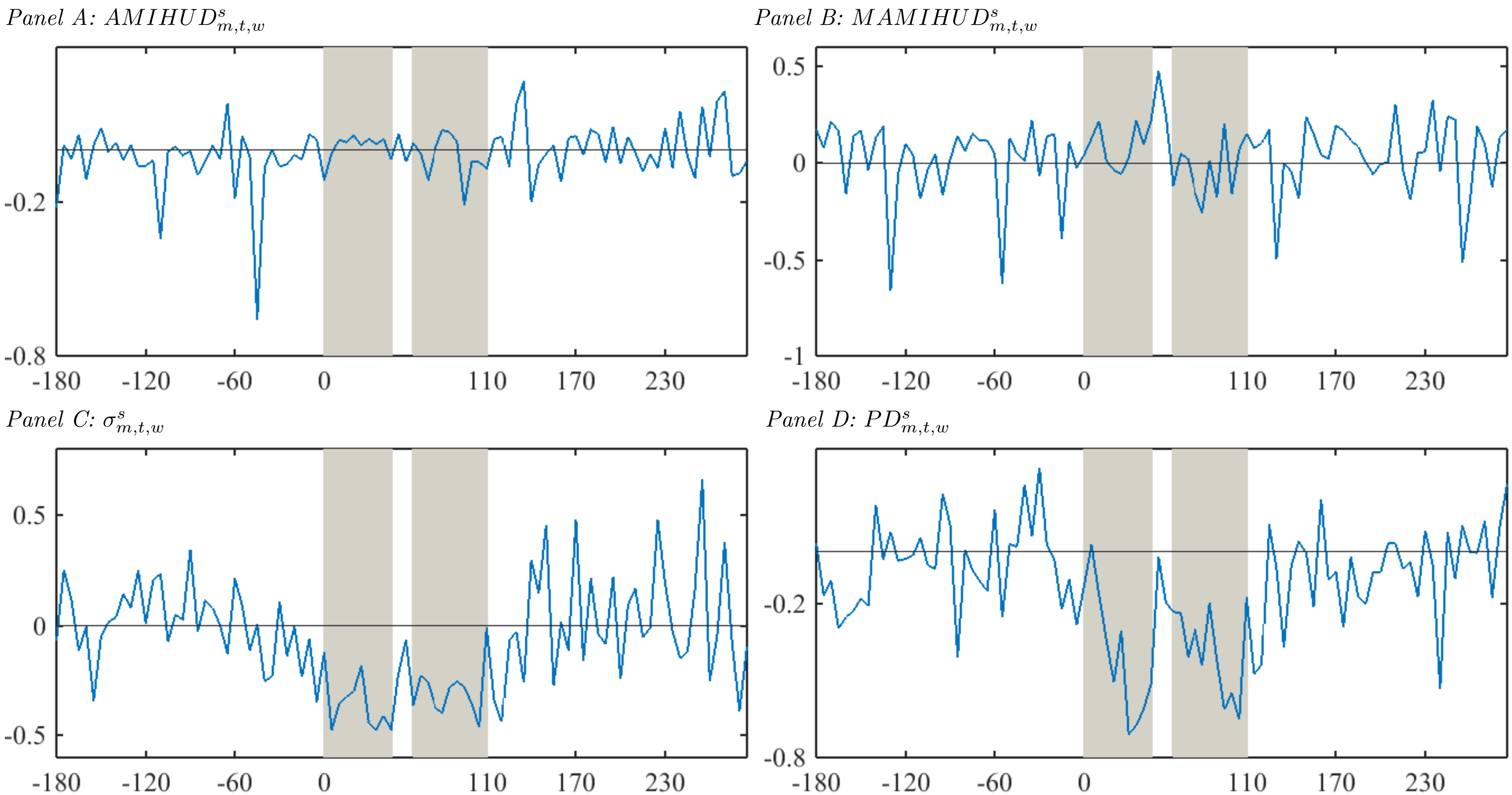


Fig. 2.5. Small Cap Market Conditions On Match Days. This figure plots the mean standardised, seasonally detrended market condition variables on match days for the small cap sample of stocks. The market variables are detrended using the methodology of Gallant et al. (1992) for month-of-the-year, day-of-the-week and time-of-the-day effects. The mean Amihud (2002) measure for market m at time t and World Cup sub-sample w is denoted by $AMIHUD^{s}_{m,t,w}$, the mean modified Amihud (2002) measure by $MAMIHUD^{s}_{m,t,w}$, mean volatility by $\sigma^{s}_{m,t,w}$ and gross price discovery by $PD^{s}_{m,t,w}$. The average first-half and second-half time periods are shaded in grey. The x-axis of each plot indicates the number of minutes from kick-off time.

Table 2.9
Small Cap Trading Activity and Market Conditions On Match Days

This table reports the estimation results for the following regression:

$$\begin{aligned} DEP_{m,t,w} =& \alpha_0 + \beta_0 GD_{m,t,w} + \sum_{\tau=1}^{6} PRE^{K-\tau}_{m,t,w}\beta_\tau + \beta_7 D^F_{m,t,w} + \beta_8 D^H_{m,t,w} + \beta_9 D^S_{m,t,w} \\ &+ \beta_{10} D^E_{m,t,w} + \sum_{\tau=1}^{6} POST^{F+\tau}_{m,t,w}\beta_{10+\tau} + \epsilon_{m,t,w} \end{aligned} \tag{2.2}$$

where $DEP_{m,t,w}$ is the dependent variable for market m at time t and World Cup sub-sample w. Each dependent variable is formed by pooling the standardised, detrended market-level observations. The $GD_{m,t,w}$ indicator variable is a match day indicator variable that takes the value of one if five-minute intra-day time period t coincides with a trading day in which country m is open for trading simultaneous to country m participating in a football match. The $D^F_{m,t,w}$ indicator variable takes the value of one if country m is playing in the first-half of a football match at time t in World Cup sub-sample w. The $D^S_{m,t,w}$ indicator variable is the analogous variable for the second-half of a football match. The $D^H_{m,t,w}$ indicator variable takes the value of one if it is half-time at time t for a match involving country m in World Cup sub-sample w. The $D^E_{m,t,w}$ indicator variable takes the value of one if it is extra-time at time t for a match involving country m in World Cup sub-sample w. The remaining indicator variables capture abnormal trading activity outside of match time on match days. The $PRE^{K-\tau}_{m,t,w}$ indicator variables take the value of one if t is 30τ minutes before a first-half kick-off observation or between 30τ and $30(\tau-1)$ minutes before a first-half kick-off observation, given that country m is participating in the match in World Cup sub-sample w. The $POST^{F+\tau}_{m,t,w}$ indicator variables take the value of one if t is 30τ minutes after a full-time observation or between 30τ and $30(\tau+1)$ minutes after a full-time observation, given that country m is participating in the match in World Cup sub-sample w. Small cap volume for market m at time t and World Cup sub-sample w is denoted by $VOL^s_{m,t,w}$, dollar volume by $DVOL^s_{m,t,w}$, number of trades by $TRADES^s_{m,t,w}$, buyer-initiated dollar trading volume by $BDVOL^s_{m,t,w}$, seller-initiated dollar trading volume by $SDVOL^s_{m,t,w}$, the mean Amihud (2002) measure by $AMIHUD^s_{m,t,w}$, the mean modified Amihud (2002) measure by $MAMIHUD^s_{m,t,w}$, the mean volatility by $\sigma^s_{m,t,w}$ and the mean gross price discovery by $PD^s_{m,t,w}$. The t-statistics are reported in italics. The standard errors are clustered at the country level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively. R^2_W is a weighted coefficient of determination that only utilises observations for which at least one non-constant independent variable is nonzero.

Panel A: Trading Activity

$DEP_{m,t,w}$	$VOL^s_{m,t,w}$	$DVOL^s_{m,t,w}$	$TRADES^s_{m,t,w}$	$BDVOL^s_{m,t,w}$	$SDVOL^s_{m,t,w}$
$GD_{m,t,w}$	-0.048	-0.082	-0.104	-0.032	-0.080*
	-1.06	*-1.61*	*-1.13*	*-0.68*	*-1.69*
$PRE^{K-6}_{m,t,w}$	0.077	0.111*	0.158*	0.138*	0.023
	1.17	*1.84*	*1.90*	*1.67*	*0.44*
$PRE^{K-5}_{m,t,w}$	0.017	0.042	0.144*	0.015	0.060
	0.34	*0.91*	*1.73*	*0.26*	*1.28*
$PRE^{K-4}_{m,t,w}$	-0.018	0.118	0.171*	0.109	0.112*
	-0.37	*1.24*	*1.84*	*1.13*	*1.65*
$PRE^{K-3}_{m,t,w}$	-0.027	0.055	0.093	0.033	0.035
	-0.38	*0.89*	*0.96*	*0.58*	*0.58*

Table 2.9
(continued)

$DEP_{m,t,w}$	$VOL^s_{m,t,w}$	$DVOL^s_{m,t,w}$	$TRADES^s_{m,t,w}$	$BDVOL^s_{m,t,w}$	$SDVOL^s_{m,t,w}$
$PRE^{K-2}_{m,t,w}$	-0.075	-0.003	-0.060	-0.007	-0.035
	-1.34	*-0.05*	*-0.51*	*-0.12*	*-0.63*
$PRE^{K-1}_{m,t,w}$	-0.001	0.017	-0.123	0.042	-0.005
	-0.02	*0.25*	*-1.01*	*0.46*	*-0.05*
$D^F_{m,t,w}$	-0.203***	-0.209**	-0.492***	-0.162***	-0.195*
	-2.63	*-2.44*	*-3.49*	*-3.01*	*-1.77*
$D^H_{m,t,w}$	-0.133*	-0.078	-0.275*	-0.028	-0.134
	-1.86	*-0.77*	*-1.73*	*-0.27*	*-1.61*
$D^S_{m,t,w}$	-0.254***	-0.199**	-0.470***	-0.172***	-0.183*
	-4.00	*-2.25*	*-3.45*	*-2.81*	*-1.89*
$POST^{F+1}_{m,t,w}$	-0.065	-0.104*	-0.183	-0.073	-0.098
	-0.89	*-1.90*	*-1.62*	*-1.33*	*-1.48*
$POST^{F+2}_{m,t,w}$	0.320*	0.088	0.170	0.043	0.139*
	1.81	*0.96*	*1.17*	*0.49*	*1.95*
$POST^{F+3}_{m,t,w}$	0.001	0.025	0.045	0.033	0.005
	0.01	*0.34*	*0.33*	*0.36*	*0.08*
$POST^{F+4}_{m,t,w}$	0.122	0.019	0.148	0.017	0.002
	1.29	*0.26*	*0.79*	*0.28*	*0.03*
$POST^{F+5}_{m,t,w}$	-0.049	0.017	0.034	0.009	-0.002
	-0.52	*0.14*	*0.24*	*0.10*	*-0.01*
$POST^{F+6}_{m,t,w}$	-0.164***	-0.083	-0.160	0.025	-0.208*
	-3.06	*-0.65*	*-0.76*	*0.15*	*-1.89*
Observations	133044	133053	133090	133052	133048
R^2_W (%)	1.82	1.81	5.19	1.12	1.69

Panel B: Market Conditions

$DEP_{m,t,w}$	$AMIHUD^s_{m,t,w}$	$MAMIHUD^s_{m,t,w}$	$\sigma_{m,t,w}$	$PD^s_{m,t,w}$
$GD_{m,t,w}$	-0.031**	-0.020	-0.035	-0.033
	-2.13	*-0.48*	*-0.57*	*-0.41*
$PRE^{K-6}_{m,t,w}$	-0.015	0.132**	0.006	-0.144
	-0.38	*2.23*	*0.13*	*-1.50*
$PRE^{K-5}_{m,t,w}$	0.033	-0.025	0.113	-0.005
	0.93	*-0.22*	*1.05*	*-0.06*
$PRE^{K-4}_{m,t,w}$	-0.045	-0.004	0.108	0.049
	-1.17	*-0.06*	*1.22*	*0.50*
$PRE^{K-3}_{m,t,w}$	0.033	0.129*	0.097	-0.078
	1.04	*1.85*	*0.93*	*-0.72*
$PRE^{K-2}_{m,t,w}$	-0.116	0.005	-0.017	0.074
	-1.43	*0.06*	*-0.17*	*0.70*
$PRE^{K-1}_{m,t,w}$	0.021	0.007	-0.080	-0.026
	0.57	*0.08*	*-1.00*	*-0.21*

Table 2.9
(continued)

$DEP_{m,t,w}$	$AMIHUD^s_{m,t,w}$	$MAMIHUD^s_{m,t,w}$	$\sigma_{m,t,w}$	$PD^s_{m,t,w}$
$D^F_{m,t,w}$	0.039	0.115*	-0.325***	-0.379*
	1.63	*1.79*	*-3.65*	*-1.93*
$D^H_{m,t,w}$	0.045	0.230	-0.177*	-0.064
	1.00	*1.48*	*-1.76*	*-0.34*
$D^S_{m,t,w}$	-0.002	0.017	-0.283***	-0.255
	-0.09	*0.19*	*-3.58*	*-1.51*
$POST^{F+1}_{m,t,w}$	0.092	-0.013	-0.176*	-0.079
	1.62	*-0.08*	*-1.69*	*-0.68*
$POST^{F+2}_{m,t,w}$	-0.014	0.076	0.135	0.116
	-0.42	*0.79*	*1.02*	*0.52*
$POST^{F+3}_{m,t,w}$	0.059*	0.096	0.117	0.045
	1.81	*1.30*	*0.81*	*0.24*
$POST^{F+4}_{m,t,w}$	0.015	0.060	0.145	0.134
	0.31	*0.69*	*0.84*	*0.84*
$POST^{F+5}_{m,t,w}$	0.047	0.058	0.132	0.069
	1.00	*0.53*	*1.06*	*0.53*
$POST^{F+6}_{m,t,w}$	0.073	0.082	-0.045	0.072
	1.51	*1.14*	*-0.25*	*0.68*
Observations	123612	126012	133081	110504
R^2_W (%)	0.46	0.29	2.43	1.89

2.6.2. Elimination-Stage Matches

This section distinguishes between ‘‘elimination-stage’’ and ‘‘group-stage’’ World Cup matches. An elimination-stage match is a match in which the losing country is immediately out of the running to win the World Cup. For this reason, elimination-stage matches are perceived to be of greater importance than group-stage matches. Of the 98 country-match observations in Table 2.1, 21 are elimination-stage matches.

As elimination-stage matches are perceived to be of greater importance than group-stage matches, the cost of monitoring the market for traders should be greater during elimination-stage matches. Therefore, it should be expected that fewer traders remain in the market during elimination-stage matches, resulting in a further reduction in trading activity during match time, relative to group-stage matches. Further, as fewer discretionary will be willing to trade during elimination-stage matches, it should be expected that the discretionary trading effect is greater for elimination-stage matches.

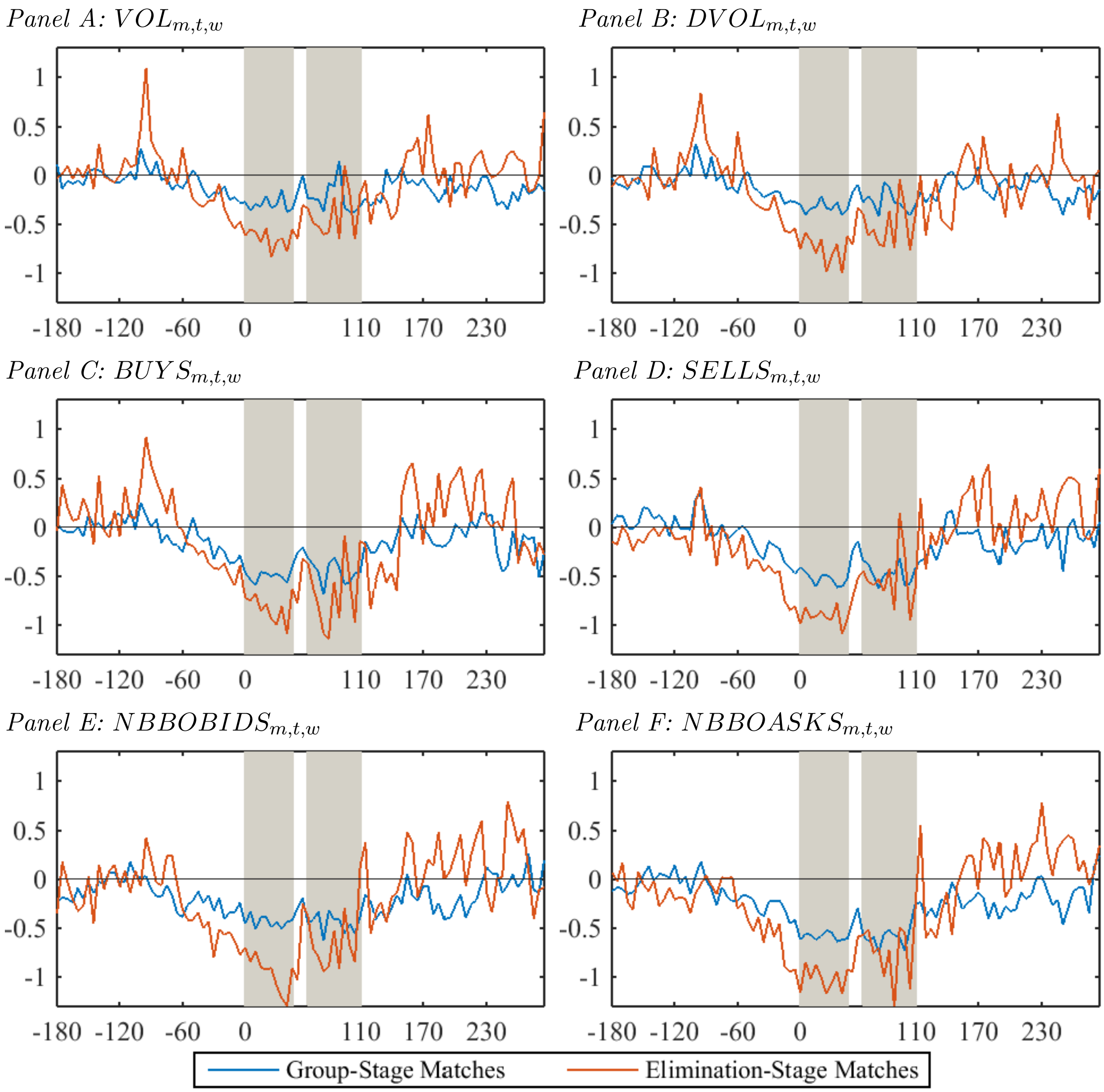


Fig. 2.6. Trading Activity On Group-Stage and Elimination-Stage Match Days. This figure plots the mean standardised, seasonally detrended trading activity variables on elimination-stage and group-stage match days. The market variables are detrended using the methodology of Gallant et al. (1992) for month-of-the-year, day-of-the-week and five-minute time-of-the-day effects. Volume for market m at time t is denoted by $VOL_{m,t,w}$, dollar volume by $DVOL_{m,t,w}$, number of trades by $TRADES_{m,t,w}$, buyer-initiated dollar trading volume by $BDVOL_{m,t,w}$, seller-initiated dollar trading volume by $SDVOL_{m,t,w}$, number of bids at the national best bid price by $NBBOBIDS_{m,t,w}$ and number of asks at the national best ask price is denoted by $NBBOASKS_{m,t,w}$. The average first-half and second-half time periods are shaded in grey. The x-axis of each plot indicates the number of minutes from kick-off time.

Figure 2.6 plots the mean $VOL_{m,t,w}$, $DVOL_{m,t,w}$, $TRADES_{m,t,w}$, $BDVOL_{m,t,w}$, $SDVOL_{m,t,w}$, $NBBOBIDS_{m,t,w}$ and $NBBOASKS_{m,t,w}$ values on elimination-stage and group-stage match-days. Several observations can be made from Figure 2.6. First, the increase in trading in the second last hour before kick-off time appears to be strongest for elimination-stage. Second, while trading activity decreases substantially during group-stage matches, trading activity falls even further during elimination-stage matches. Thus, Figure 2.6, provides strong evidence that football matches are the cause of the abnormal trading activity that occurs on match days and that elimination-stage matches are associated with greater monitoring costs for traders than group-stage matches.

2.7. Conclusion

This paper documents a unique consequence of limited attention in financial markets. This paper finds evidence of discretionary trading whereby market participants trade before World Cup football matches in order to avoid trading during matches. This effect results in an abnormal amount of trading in anticipation of World Cup football matches, particularly from 120 to 90 minutes until kick-off time. The discretionary trading effect combined with the contemporaneous distraction effect of World Cup football matches result in exceptional market conditions during match days. The resulting market trends do not conform to the intra-day market trends previously documented in the literature. The empirical analysis indicates that the pre-match period of high discretionary trading is accompanied by reduced price impact costs, increased volatility and increased price discovery. During match time markets exhibit increased price impact costs, reduced volatility and reduced price discovery. The extraordinary market conditions observed during these unique trading days adhere to the theoretical predictions of the Admati and Pfleiderer (1988) model of discretionary trading.

This paper has a number of important implications. First, this paper demonstrates that the impact of distraction events on markets is more complex and significant than has previously been documented. When culturally significant market-orthogonal events can be anticipated, there is a contemporaneous and asynchronous reaction by market participants. Second, the dual response of market participants to anticipated distraction events creates a unique landscape of market conditions that is suited to dynamic trading strategies. Uninformed traders can benefit from the reduced price impact costs that occur in anticipation of distraction events; while, informed traders can gain greater profits from price inefficiencies in this period. Uninformed traders should avoid trading during distraction events as they will incur greater transaction costs; while, predatory informed traders could benefit from the reduced monitoring that occurs during distraction periods. Third, in contrast to the traditional Admati and Pfleiderer (1988) discretionary trader, this paper demonstrates that shocks to the opportunity cost of monitoring the market can also drive discretionary trading. In conjunction with the pooling Nash equilibria of the Admati and Pfleiderer (1988) model, this means that even minor distraction events have the potential to significantly distort markets.

3. Sports Sentiment and Stock Returns: An Intra-day Study

3.1. Abstract

In their influential paper, Edmans et al. (2007) demonstrate that sporting results can predict overnight stock returns. The authors attribute this to a sports sentiment effect. I demonstrate that the Edmans et al. (2007) daily sentiment effect is still present in a more recent sample of stock market data. In addition, I utilise all FIFA World Cup matches that have occurred during trading hours since 1998 to determine that there is an analogous intra-day sentiment effect. Winning full-time outcomes are associated with positive abnormal stock returns for the remainder of the trading day. Moreover, unexpected victories and victories over traditional rivals have a significant and positive marginal impact on abnormal stock returns. Using trade and quote data, this study also documents abnormal order imbalance and quote revision activity surrounding half-time match outcomes. Evidence suggests that both liquidity takers and providers are influenced by investor sentiment. Small trades exhibit the greatest sentiment effects.

JEL Classifications: G11, G12, G14, G41, L83

Keywords: Stock Returns, Investor Sentiment, FIFA World Cup

3.2. Introduction

This paper provides evidence of the intra-day impacts of investor mood on trading behaviour and confirms previous findings of the overnight impacts of investor sentiment on abnormal stock returns. Following previous studies, this paper takes advantage of sporting outcomes to identify exogenous shocks to investor sentiment. Consistent with the previous literature, I find that losses in FIFA World Cup matches are associated with a -17.1 basis point overnight abnormal stock return for the losing country. Using a smaller sample of FIFA World Cup football matches that coincide with the trading hours of participating countries, I test for analogous sentiment effects relating half-time and full-time match outcomes. There is evidence of a win-effect whereby immediately after a football match concludes, the winning country on average experiences 0.5-0.6 basis point abnormal stock returns every five minutes between full-time and the close of trading. Further, unexpected wins and wins over traditional football rivals have a significant positive marginal impact on intra-day stock returns. The final portion of the paper utilises trade and quote data to directly test whether sentiment impacts on investors' buying and selling behaviour, as well as liquidity providers' quote-setting behaviour. The empirical analysis reveals that order imbalance measures are correlated to half-time match outcomes. After a losing half-time match outcome most order imbalance measures decrease by at least 25% of a standard deviation. The order imbalance results are stronger for small trades. This suggests that small trades are more susceptible to sentiment effects. Finally, there is some evidence that non-trade-driven quote revisions are similarly-impacted.

This paper is part of a growing literature that examines the impact of investor sentiment on financial decision-making. The greatest difficulty associated with this area of research is the nature of investor sentiment itself. Investor sentiment is unobservable and non-quantifiable. Accordingly, a number of approaches have been taken to overcome these difficulties. One approach has been to identify sudden shifts in sentiment. For example, Kamstra, Kramer, and Levi (2000) associate disruptions to sleeping patterns with negative sentiment shocks. Kamstra et al. (2000) use daylight saving to identify shifts in sleeping patterns.[18] Another innovative strategy for identifying shifts to sentiment is proposed by Ashton et al. (2003). Ashton et al. (2003) propose utilising the outcomes of international football matches as an instrument for sentiment shocks. This novel approach has a number of advantages. First, football has a very large global following.[19] Second, the fortunes of national football teams can easily be mapped to national stock markets and market indices. Third, the outcomes of individual football matches are largely removed from stock market fundamentals.

In a study most closely related to this one, Edmans et al. (2007) extend on the seminal findings of Ashton et al. (2003) by providing a comprehensive and global study of sports sentiment.[20] Edmans et al. (2007) find very strong evidence of a loss-effect whereby losses in sporting matches are associated with negative abnormal daily stock returns. For example, a loss in a FIFA Football World Cup (henceforth "World Cup") Elimination Stage match is associated with a -49 basis point abnormal daily return. Edmans et al. (2007) demonstrate that the sports sentiment effect is strongest for World Cup football matches. In a concurrent study, Cai, Fan, Ko, Richione, and Russo (2018) attribute the overnight loss-effect documented by Edmans et al. (2007) to both physiological and psychological factors. Cai et al. (2018) argue that a proportion of the overnight loss-effect can be attributed to the physiological impacts of sleeplessness. The sleeplessness effect is due to investors disrupting their regular sleeping patterns to monitor football matches. In this sense, the arguments of Cai et al. (2018) are related to those presented in Kamstra et al. (2000). Accordingly, the physiological impacts of sleeplessness may counteract positive overnight sentiment shocks, resulting in the asymmetric loss-effect documented by Edmans et al. (2007). One should not expect a similar dampening of positive sentiment to occur for winning football outcomes that occur during regular trading hours. Thus, the empirical results of this paper lend support to Cai et al. (2018), who distinguish between physiological and psychological impacts on trading behaviour.

To date, the vast majority of studies within the sports sentiment literature have concentrated on stock market performance at the daily frequency. This is because most sporting events are held

[18] Kamstra et al. (2000) demonstrate that daylight saving weekend returns are 200-500% more negative than ordinary weekend returns. Pinegar (2002) argues that the Kamstra et al. (2000) study is compromised by outlier observations. Kamstra, Kramer, and Levi (2002) question the econometric methodologies employed by Pinegar (2002).

[19] For example, according to FIFA's 2010 World Cup Television Audience Report (http://www.fifa.com/mm/document/affederation/tv/01/47/32/73/2010fifaworldcupsouthafricatvaudiencereport.pdf) 3.2 billion people watched at least one minute of the 2010 World Cup, while 909.6 million people watched at least one minute of the 2010 World Cup final.

[20] Other studies of sports sentiment include: Kaplanski and Levy (2010);Mishra and Smyth (2010);Chang et al. (2012);Kaplanski and Levy (2014);Pantzalis and Park (2014); and, Kaplanski, Levy, Veld, and Veld-Merkoulova (2015).

during the evening hours, when television audiences are their largest and equity markets are closed. Thus, the timing of sporting events restricts the ability to conduct an intra-day study of sports sentiment. Nonetheless, in their innovative studies, Ehrmann and Jansen (2016, 2017) highlight the unique implications of the organisational structure of the World Cup. Each World Cup generally takes place in one "host" country.[21] Further, all World Cup matches take place in the afternoon and evening hours of the host country. This means that, on occasion, some matches take place during the trading hours of the participating countries. As a result, the timing of football matches, relative to stock market trading hours, is determined in a quasi-random fashion. This provides a unique opportunity to study the intra-day effects of investor sentiment on financial markets.

At the time of writing and to the author's knowledge, Ehrmann and Jansen (2016) is the only study to examine *intra-day* sports sentiment effects. Ehrmann and Jansen (2016) provide evidence of underpricing due to losing football match outcomes. Unfortunately, the Ehrmann and Jansen (2016) study is restricted by a modest sample size. The Ehrmann and Jansen (2016) study only considers one cross-listed stock and two shocks to investor sentiment (two football matches). In contrast to Ehrmann and Jansen (2016) and in the spirit of Edmans et al. (2007), the objective of this paper is to provide a comprehensive and global examination of intra-day sentiment effects. This is achieved by considering all World Cup football matches that have occurred during trading hours since 1998.[22] The intra-day analysis presented in this study encompasses 19 equity markets and 106 match outcomes. Of the 106 match outcomes, 53 are winning or losing half-time match outcomes. The remainder are winning or losing full-time match outcomes that occurred during trading hours.

This sample of intra-day shocks to investor sentiment allows for a unique contribution to the sentiment literature. The conjecture of the sentiment literature is that sentiment influences investors' levels of optimism (or pessimism). These shifts in optimism, in turn, influence investors' likelihood of buying or selling shares. Previous studies *indirectly* test this argument by observing abnormal stock returns. Previous studies do not directly test their hypothesis by observing abnormal order imbalances. In this study, I utilise the Lee and Ready (1991) algorithm to classify trades as buyer-initiated or seller-initiated. In support of the sentiment hypothesis, the empirical analysis presented in this paper reveals abnormal buying and selling behaviour can be predicted by half-time match outcomes that occur during trading hours. This suggests that sentiment does in fact influence trading behaviour. Further analysis dissects trades into dollar volume quintiles and constructs five corresponding dollar order imbalance measures. The abnormal order imbalance results are stronger for small trades. This suggests that small trades are more susceptible to sentiment shocks.

One might also consider the possibility that liquidity providers are influenced by sentiment. If this is the case, liquidity providers might revise their quotes upwards following a positive sentiment

[21] The one exception is the 2002 World cup. The 2002 World Cup was hosted by both the Korea Republic and Japan.

[22] This date is chosen on the basis of data availability.

shock and revise their quotes downwards following a negative sentiment shock. This paper presents empirical evidence of downward bid-quote revisions and no significant revision to ask-quotes following negative sentiment shocks, particularly after half-time loss outcomes. This suggests that sentiment-induced selling causes liquidity providers to revise mid-quotes downwards while simultaneously increasing spreads. Thus, liquidity providers appear to interpret sentiment-induced abnormal order flow as possibly informed and accordingly adjust spreads to compensate for adverse selection costs. Thus, liquidity providers appear to be indirectly influenced by investor sentiment.

The remainder of this paper is structured as follows. Section 3.3 describes the data utilised in the study. Section 3.4 presents the empirical results. Concluding remarks are presented in Section 3.5.

3.3. Data and Summary Statistics

This study makes use of four datasets. The first dataset contains information regarding all World Cup football matches between 1974 and 2014. The remaining datasets contain stock market information of increasing granularity. The second dataset is comprised of daily stock market returns. The third dataset is comprised of five-minute intra-day stock market returns. The final dataset contains trade and quote information.

3.3.1. Football Match Data

This study utilises World Cup football match outcomes to identify shocks to investor sentiment. This methodology has a number of advantages. Firstly, individual football matches are largely removed from asset fundamentals. This allows for a relatively clean test of the impact of sentiment on investor behaviour. Second, World Cup football matches are scheduled in a quasi-random fashion relative to the domestic trading hours of the participating countries. This is because all matches are hosted in one country (or geographical region) and scheduled to coincide with the afternoon and evening hours of the "host" country (or region). Thus, if a participating country's domestic trading hours coincide with the afternoon or evening hours of the host country, the participating country's main exchange may be open for trading while their national football team plays in a World Cup football match. This allows for a very unique analysis of investor sentiment at the intra-day level. The third advantage is that World Cup football match outcomes are time-stamped to a high degree of accuracy. This is not true of other sentiment shocks that are often observed at the daily frequency. The fourth advantage is the wide appeal of the World Cup. The large following of the World Cup allows for a comprehensive and global study of investor sentiment. The most significant disadvantage of utilising World Cup football matches to identify shocks to investor sentiment is the modest number of observations. This is because the World Cup is only held every four years. Nonetheless, World Cup football matches have been used as an identification tool in a number of studies of investor sentiment.[23]

Following Ehrmann and Jansen (2017), the football match data is extracted from the official World Cup Match Reports available at the FIFA online archive.[24] I extract data regarding all World Cup football matches from 1974 to 2014. During this time period, there were 11 iterations of the World Cup. As in Edmans et al. (2007), World Cups prior to 1974 are not considered. This is due to a lack of international daily stock market return observations prior to 1973. For each match, I extract the full-time score, the half-time score, the match date, the match location, the scheduled kick-off time, the amount of extra stoppage time that took place at half-time or full-time and the amount of extra-time, if any. Using this information, I can estimate the time at which the first-half and the entire match concluded. These estimates are not entirely accurate as matches do not necessarily begin as scheduled. Further, matches may not recommence exactly as scheduled after half-time. That is, half-time observations may not always extend for exactly 15 minutes. Nonetheless, I assume that matches begin as scheduled and that every half-time observation extends for exactly 15 minutes to arrive at my half-time and full-time time-stamps.

3.3.2. Daily Stock Market Returns

The first stock market dataset used in this study is a dataset of daily stock market returns. This dataset is constructed in a similar fashion to Edmans et al. (2007). The daily stock market returns are extracted from Datastream. I extract daily stock returns for 41 countries. This set of countries is determined by data availability and participation in at least one World Cup since 1974. Following Edmans et al. (2007), where possible, I utilise total return indices for each country to arrive at stock market returns. The total return indices assume that dividends are reinvested. Following Edmans et al. (2007), the variable of interest for each market is the daily continuously compounded return.

The 41 countries represented in the daily stock market return analysis differ from those in the Edmans et al. (2007). There are several reasons for this difference. First, Datastream now provides greater market coverage. Second, more countries have participated in World Cups since the original Edmans et al. (2007) study. Third, Edmans et al. (2007) choose to exclude the United States and Canada from their analysis. Their justification is that football is not a prominent sport in these countries.

The daily stock market returns data is summarised in Table 3.1. The starting date for each country is dictated by data availability. The final observation for every stock market is the 30th of December 2014. Table 3.1 also summarises the World Cup match outcomes over each country's respective sample period. For example, the total market index of Argentina is TOTMKAR. TOTMKAR data is available from the 8th of January 1990 until the end of the sample period,

[23] These studies include Ashton et al. (2003); Edmans et al. (2007); Kaplanski and Levy (2010, 2014); Ehrmann and Jansen (2016) and Cai et al. (2018).

[24] http://www.fifa.com/fifa-tournaments/archive/worldcup/index.html.

Table 3.1
Daily Stock Market Returns Data
This table outlines the daily stock market returns data extracted from Datastream. This table also summarises the FIFA World Cup match outcomes of each country over the country's respective sample period. Note that Germany refers to the unified Germany and West Germany.

Country	Datastream Mnemonic	Start Date	End Date	Wins	Draws	Losses
Argentina	TOTMKAR	08-Jan-1990	31-Dec-2014	15	3	4
Australia	TOTMKAU	09-Jan-1973	31-Dec-2014	2	1	5
Austria	TOTMKOE	27-Apr-1983	30-Dec-2014	1	2	1
Belgium	TOTMKBG	09-Jan-1973	31-Dec-2014	11	5	8
Brazil	TOTMKBR	11-Jul-1994	30-Dec-2014	18	1	3
Canada	TOTMKCN	09-Jan-1973	31-Dec-2014	0	0	2
Chile	TOTMKCL	11-Jul-1989	31-Dec-2014	3	3	2
China	TOTMKCH	27-Jul-1993	31-Dec-2014	0	0	2
Colombia	TOTMKCB	16-Jan-1992	30-Dec-2014	4	0	2
Croatia	TOTMKCT	04-Oct-2005	31-Dec-2014	0	2	3
Czech Republic	TOTMKCZ	15-Mar-1994	30-Dec-2014	1	0	1
Côte d'Ivoire	IFGECD$	17-Oct-2008	31-Oct-2014	0	1	3
Denmark	TOTMKDK	08-Jan-1982	31-Dec-2014	4	2	4
Ecuador	IFFMECL	01-Aug-2008	30-Dec-2014	0	1	1
England	TOTMKUK	02-Jan-1973	31-Dec-2014	11	7	7
France	TOTMKFR	09-Jan-1973	31-Dec-2014	14	6	8
Germany	TOTMKBD	09-Jan-1973	30-Dec-2014	32	9	7
Ghana	IFFMGHL	01-Aug-2008	31-Dec-2014	1	0	3
Greece	TOTMKGR	12-Jan-1988	31-Dec-2014	2	1	5
Italy	TOTMKIT	09-Jan-1973	30-Dec-2014	21	10	8
Japan	TOTMKJP	10-Jan-1973	30-Dec-2014	3	3	6
Mexico	TOTMKMX	15-Apr-1988	31-Dec-2014	5	4	8
Morocco	TOTMKMC	02-Jan-1996	31-Dec-2014	1	1	1
Netherlands	TOTMKNL	09-Jan-1973	31-Dec-2014	18	7	8
Nigeria	TOTMKNG	09-Sep-2009	31-Dec-2014	0	2	3
Norway	TOTMKNW	21-Feb-1980	30-Dec-2014	2	3	1
Poland	TOTMKPO	08-Mar-1994	30-Dec-2014	1	0	2
Portugal	TOTMKPT	23-Jan-1990	31-Dec-2014	6	2	4
Republic of Ireland	TOTMKIR	11-Jan-1978	31-Dec-2014	2	5	2
Republic of Korea	TOTMKKO	16-Sep-1987	30-Dec-2014	3	6	8
Romania	TOTMKRM	09-May-1997	31-Dec-2014	2	0	1
Russia	TOTMKRS	28-Jan-1998	31-Dec-2014	1	2	2
Saudi Arabia	IFGDSBL	02-Jan-1998	31-Dec-2014	0	1	2
Slovakia	TOTMKSX	02-Mar-2006	31-Dec-2014	1	1	2
Slovenia	TOTMKSJ	04-Jan-1999	30-Dec-2014	1	0	3
South Africa	TOTMKSA	09-Jan-1973	31-Dec-2014	1	3	2
Spain	TOTMKES	09-Mar-1987	31-Dec-2014	14	2	4
Sweden	TOTMKSD	12-Jan-1982	30-Dec-2014	2	6	4
Switzerland	TOTMKSW	02-Feb-1973	30-Dec-2014	5	1	4
Turkey	TOTMKTK	12-Jan-1988	31-Dec-2014	2	1	2
United States of America	TOTMKUS	03-Jan-1973	31-Dec-2014	5	2	12

Table 3.2
Daily Stock Market Returns Summary Statistics
This table provides the summary statistics for each series of daily stock market returns.

Country	Mean	Median	Standard Deviation	Skewness	Kurtosis
Argentina	0.087	0.012	2.341	-1.673	85.862
Australia	0.027	0.042	1.087	-2.173	59.743
Austria	0.032	0.051	1.103	-0.382	12.330
Belgium	0.024	0.039	0.969	-0.343	13.279
Brazil	0.043	0.078	1.635	0.094	12.635
Canada	0.026	0.055	0.921	-0.765	16.432
Chile	0.053	0.039	0.929	0.104	8.378
China	0.031	0.012	1.968	0.116	9.429
Colombia	0.050	0.030	1.039	-0.202	16.507
Croatia	0.010	0.021	1.188	0.268	18.430
Czech Republic	0.007	0.035	1.332	-0.373	15.178
Denmark	0.041	0.053	1.052	-0.466	10.736
Ecuador	-0.012	-0.001	1.189	-0.789	75.215
England	0.028	0.052	1.088	-0.229	11.046
France	0.029	0.044	1.196	-0.250	8.229
Germany	0.021	0.056	1.074	-0.218	14.679
Ghana	0.072	0.051	0.935	-0.576	14.612
Greece	0.017	0.015	1.829	-0.071	8.870
Republic of Ireland	0.032	0.045	1.203	-0.765	14.533
Italy	0.027	0.042	1.374	-0.228	7.574
Côte d'Ivoire	0.025	-0.004	8.432	-0.731	719.957
Japan	0.014	0.020	1.145	-0.405	13.698
Republic of Korea	0.024	0.018	1.829	0.018	7.504
Mexico	0.079	0.060	1.346	0.090	9.879
Morocco	0.030	0.023	0.735	-0.144	14.097
Netherlands	0.022	0.049	1.103	-0.330	10.561
Nigeria	0.051	0.018	0.992	0.288	11.334
Norway	0.033	0.055	1.461	-0.660	14.080
Poland	0.009	0.041	1.731	-0.222	9.226
Portugal	0.002	0.023	1.045	-0.311	11.551
Romania	0.031	0.039	2.018	-0.462	15.680
Russia	0.076	0.053	2.665	0.184	15.938
Saudi Arabia	0.042	0.104	1.726	-0.856	15.338
Slovakia	-0.004	0.015	0.935	-3.351	84.605
Slovenia	0.012	0.022	0.965	-0.431	11.764
South Africa	0.053	0.064	1.290	-0.625	11.632
Spain	0.022	0.067	1.268	-0.160	8.448
Sweden	0.044	0.071	1.403	-0.021	7.670
Switzerland	0.023	0.047	0.944	-0.702	15.763
Turkey	0.135	0.040	2.495	-0.005	7.596
United States of America	0.028	0.052	1.081	-0.864	23.162

Table 3.3
Five-minute Intra-day Stock Market Returns Data
This table outlines the five-minute intra-day stock market returns data extracted from Thomson Reuters Tick History.

Country	Market Index	Start Date	End Date
Argentina	MERV	02-Jan-1998	30-Dec-2014
Belgium	BFX	02-Jan-1998	31-Dec-2014
Brazil	BVSP	02-Jan-1998	30-Dec-2014
Chile	IPSA	02-Jan-1998	30-Dec-2014
Denmark	OMXC20	02-Jan-2006	30-Dec-2014
England	FTSE	02-Jan-1998	31-Dec-2014
France	FCHI	02-Jan-1998	31-Dec-2014
Germany	GDAXI	02-Jan-1998	30-Dec-2014
Ireland	ISEQ	02-Jan-1998	31-Dec-2014
Italy	SPMIB	02-Jan-1998	29-May-09
	FTMIB	1-Jun-09	31-Dec-2014
Mexico	MXX	02-Jan-1998	31-Dec-2014
Netherlands	AEX	02-Jan-1998	31-Dec-2014
Poland	WIG20	02-Jan-1998	31-Dec-2014
Portugal	PSI20	02-Jan-1998	31-Dec-2014
South Africa	JDTOP	02-Jan-1998	31-Dec-2014
Spain	IBEX	02-Jan-1998	31-Dec-2014
Switzerland	SSMI	02-Jan-1998	31-Dec-2014
Turkey	XU030	02-Jan-1998	31-Dec-2014
United States of America	DJI	02-Jan-1998	31-Dec-2014

the 30th of December 2014. During that time, Argentina was involved in 22 World Cup football matches. Fifteen of those matches were wins for Argentina, three were draws and four were losses.

Table 3.2 provides the summary statistics of the daily stock market return series. Of the 41 countries in the daily stock return sample, 39 have a positive mean daily return. Ecuador and Slovakia have a negative mean daily return over their respective sample periods. The standard deviation of daily stock market returns varies over the cross-section. The country with the lowest standard deviation is Morocco with a standard deviation of 0.735 and the country with the largest standard deviation is Côte d'Ivoire with a standard deviation of 8.432. Table 3.2 reveals that the majority of countries have negatively skewed returns. Further, every country has a kurtosis value greater than three, indicating a leptokurtic distribution.

3.3.3. Five-minute Intra-day Stock Market Returns

The second stock market dataset used in this study is a dataset of five-minute intra-day stock market returns. The data is extracted from Thomson Reuters Tick History.[25] The dataset is comprised of intra-day return series from 19 countries. To arrive at the return series, I first nominate a national stock market index for each country. Following, I take the five-minute intra-day return

[25]Access to Thomson Reuters Tick History was provided by the Securities Industry Research Centre of Asia-Pacific (SIRCA).

Table 3.4
Five-minute Intra-day Stock Market Returns Summary Statistics
This table provides the summary statistics for each series of five-minute intra-day stock market returns.

Country	Mean	Median	Standard Deviation	Skewness	Kurtosis
Argentina	-0.152%	0.000%	0.135	0.270	65.555
Belgium	-0.043%	0.000%	0.095	-0.382	40.443
Brazil	-0.034%	0.000%	0.148	0.022	207.667
Chile	-0.031%	0.000%	0.069	2.087	1101.554
Denmark	-0.066%	0.000%	0.101	-0.206	365.275
England	-0.027%	0.000%	0.096	-1.418	227.339
France	-0.021%	0.000%	0.118	-0.118	21.052
Germany	-0.018%	0.000%	0.128	-0.255	24.160
Ireland	-0.045%	0.000%	0.159	-0.375	932.784
Italy	-0.054%	0.000%	0.125	0.684	209.905
Mexico	0.003%	0.000%	0.094	0.061	58.057
Netherlands	-0.030%	0.000%	0.111	-0.180	31.013
Poland	-0.007%	0.000%	0.133	0.383	30.559
Portugal	-0.063%	0.000%	0.095	0.009	39.845
South Africa	0.021%	0.000%	0.094	0.055	19.306
Spain	-0.002%	0.086%	0.123	0.042	24.398
Switzerland	-0.037%	0.000%	0.095	-0.092	20.440
Turkey	0.033%	0.180%	0.315	-0.575	69.442
United States of America	0.009%	0.000%	0.109	0.016	29.803

of each index as my variable of interest. For a country to be included in the five-minute intra-day stock market returns dataset, the country's national stock market must be open for trading during at least one half-time or full-time match observation involving the national team associated with the country. For each country, the sample period is from January 1998 to December 2014.[26] All countries have a continuous series of stock returns for the entire sample period.[27] Table 3.3 gives the market index nominated for each country and the sample period for each country in the five-minute intra-day stock market return dataset.

Table 3.4 gives the summary statistics for the five-minute intra-day stock market returns. Notably, most countries have a negative mean five-minute intra-day stock market return. This is consistent with the recent findings of Berkman, Koch, Tuttle, and Zhang (2012). Berkman et al. (2012) use US stock market data from 1996 to 2008 to show that overnight stock returns tend to be positive; while, intra-day stock market returns tend to be negative. Table 3.4 also demonstrates that most markets have a median five minute intra-day return of 0. This could be attributed to periods of zero price-impact. Some markets exhibit positively skewed intra-day return distributions, while others exhibit negatively skewed distributions. Finally, all intra-day return distributions are highly leptokurtic. This is consistent with the daily stock return distributions.

[26]Intra-day stock market returns from Thomson Reuters Tick History become available from 1996. The first iteration of the World Cup following 1996 is the 1998 iteration. Thus, the January 1998 to December 2014 sample period is designed to capture all World Cup iterations since 1996.

[27]Note that in 2009, the Borsa Italiana appointed FTSE Group as the new provider of the MIB index, taking over from Standard and Poor's.

Table 3.5
Summary of Match Outcomes that Occurred during Trading Hours
This table summarises the match outcomes that occurred during trading hours for each country. HW signifies a half-time win outcome. HL signifies a half-time loss outcome. GHW signifies a Group Stage half-time win outcome. GHL signifies a Group Stage half-time loss outcome. EHW signifies an Elimination Stage half-time win outcome. EHL signifies an Elimination Stage half-time loss outcome. W signifies a full-time win outcome. L signifies a full-time loss outcome. GW signifies a Group Stage full-time win outcome. GL signifies a Group Stage full-time loss outcome. EW signifies an Elimination Stage full-time win outcome. EL signifies an Elimination Stage full-time loss outcome.

Country	HW	HL	GHW	GHL	EHW	EHL	W	L	GW	GL	EW	EL
Argentina	2	0	2	0	0	0	3	0	2	0	1	0
Belgium	1	0	1	0	0	0	1	1	1	0	0	1
Brazil	4	1	1	0	3	1	5	2	3	0	2	2
Chile	2	1	2	1	0	0	1	1	1	1	0	0
Denmark	0	0	0	0	0	0	0	1	0	1	0	0
England	3	0	3	0	0	0	3	1	3	0	0	1
France	0	3	0	3	0	0	0	3	0	3	0	0
Germany	3	1	2	1	1	0	3	1	1	1	2	0
Ireland	1	1	1	1	0	0	1	0	1	0	0	0
Italy	3	1	2	1	1	0	1	0	1	0	0	0
Mexico	0	3	0	3	0	0	1	3	1	2	0	1
Netherlands	1	1	0	0	1	1	3	0	1	0	2	0
Poland	1	2	1	2	0	0	1	2	1	2	0	0
Portugal	3	2	3	2	0	0	2	2	2	2	0	0
South Africa	1	0	1	0	0	0	0	0	0	0	0	0
Spain	3	1	3	1	0	0	3	0	3	0	0	0
Switzerland	1	0	1	0	0	0	1	0	1	0	0	0
Turkey	2	0	1	0	1	0	2	1	1	0	1	1
United States of America	0	5	0	5	0	0	1	3	1	3	0	0
Total	31	22	24	20	7	2	32	21	24	15	8	6

Table 3.5 gives a summary of the match outcomes that occurred during the trading hours of each country represented in the five-minute intra-day stock market return dataset. During the sample period, there are 31 winning half-time observations and 22 losing half-time outcomes. Further, there are 32 winning full-time outcomes and 21 losing full-time outcomes. Thus, there are a total of 106 intra-day match outcome variables.

Each iteration of the World Cup is separated into an initial "Group Stage" and a latter "Elimination Stage". Elimination Stage matches are perceived to be of greater importance as the loser of an Elimination Stage match is immediately eliminated from the remainder of the World Cup. During the sample period, there are 24 Group Stage winning half-time match outcomes, 20 Group Stage losing half-time match outcomes, 24 winning Group Stage full-time match outcomes and 15 losing Group Stage full-time match outcomes. Unfortunately, there are few match outcomes relating to Elimination Stage matches in the sample. This is because for each World Cup, the majority of matches are Group Stage matches. During the sample period, there are seven Elimination Stage winning half-time match outcomes, two Elimination Stage losing half-time match outcomes, eight Elimination Stage winning full-time match outcomes and six Elimination Stage losing full-time

Table 3.6
Market Indices
This table details the market index sample. The representative stocks for each country are sampled from the following national indices.

Country	1998	2002	2006	2010	2014
Argentina	MERV		MERV	MERV	MERV
Belgium	BFX	BFX			
Brazil	BVSP		BVSP	BVSP	
Chile				IPSA	IPSA
Denmark	KFX			OMXC20	
England	FTSE	FTSE		FTSE	
France	FCHI	FCHI		FCHI	
Germany	GDAXI	GDAXI	GDAXI	GDAXI	
Ireland		ISEQ			
Italy			SPMIB	FTMIB	
Mexico	MXX		MXX	MXX	MXX
Netherlands	AEX			AEX	
Poland		WIG20	WIG20		
Portugal		PSI20	PSI20	PSI20	
South Africa				JDTOP	
Spain		IBEX	IBEX	IBEX	
Switzerland			SSMI	SSMI	
Turkey		XU030			
United States of America	DJI		DJI	DJI	DJI

match outcomes.

3.3.4. Trade and Quote Data

The third stock market dataset used in this study is comprised of trade and quote observations. The trade and quote data is extracted from Thomson Reuters Tick History. Similar to the intra-day stock market returns data, the trade and quote data becomes available from 1996. At the time of writing, there have been five iterations of the World Cup since 1996. Accordingly, the trade and quote dataset is comprised of five sub-samples. Each sub-sample corresponds to a World Cup iteration. Since 1996, the majority of World Cup games have occurred during the month of June. Thus, for each World Cup, I extract trade and quote data from the months of May, June and July. This sampling technique is designed to reduce computational difficulty and include a significant amount of time before and after each World Cup. For each country and World Cup iteration, I extract trade and quote data from the constituents of a nominated national market index. The nominated market indices for each market and World Cup iteration are presented in Table 3.6.

The countries represented in the trade and quote dataset vary across the World Cup sub-samples. This is because the individual World Cup sub-sample countries are independently determined. That is, for country m to be included in World Cup sub-sample w, the country's national stock market must be open for trading during at least one half-time or full-time match observation

Table 3.7
Stock Market Trading Hours
This table details the normal trading hours of each national stock market during each iteration of the FIFA World cup from 1998 to 2014.

Country	1998	2002	2006	2010	2014
Argentina	11:00-17:00		11:00-17:00	11:00-17:00	11:00-17:00
Belgium	10:00-16:30	9:00-17:20			
Brazil	10:00-17:00		10:00-17:00	10:00-17:00	
Chile				9:30-16:00	9:30-16:00
Denmark	9:00-17:00			9:00-17:00	
England	8:30-16:30 (1/5-19/7) 9:00-16:30 (20/7-31/7)	8:00-16:30		8:00-16:30	
France	10:00-17:00	9:00-17:30		9:00-17:30	
Germany	8:30-17:00	9:00-20:00	9:00-17:30	9:00-17:30	
Ireland		8:00-17:30			
Italy			9:00-17:25	9:00-17:25	
Mexico	8:30-15:00		8:30-15:00	8:30-15:00	8:30-15:00
Netherlands	9:30-16:30			9:00-17:30	
Poland		10:00-16:00	10:00-16:00		
Portugal		8:00-16:30	8:00-16:30	8:00-16:30	
South Africa				9:30-17:00	
Spain		9:00-17:30	9:00-17:30	9:00-17:30	
Switzerland			9:00-17:20	9:00-17:20	
Turkey		9:30-12:00 & 14:00-16:30			
United States of America	9:30-16:00		9:30-16:00	9:30-16:00	9:30-16:00

during World Cup m, conditional on the match involving the national team associated with country m. For example, Table 3.6 indicates that in 2014, there were four countries that were open for trading during at least one match outcome: Argentina; Chile; Mexico and the United States of America. European countries are not represented in the 2014 sub-sample because European exchanges were closed for trading during the afternoon and evening hours of the host country, Brazil. In contrast to the 2014 sub-sample, the 2010 sub-sample includes 15 countries. This is because the host country of the 2010 World Cup was South Africa. Many European and American exchanges were open for trading during the afternoon and evening hours of South Africa. Thus, the domestic trading hours of each country are crucial for observing the impacts of intra-day match observations. Accordingly, Table 3.7 gives the trading hours of each country in the trade and quote dataset, for each World Cup iteration. Similar to the five-minute intra-day sample, the trade and quote sample comprises of data from 19 countries across all iterations of the World Cup.

3.4. Empirical Analysis

The empirical analysis is divided into three parts. Each part corresponds to a different dataset described in Section 3.3. The first portion of the empirical analysis utilises the daily stock market returns dataset described in Sub-section 3.3.2. The objective of this preliminary analysis is to confirm the findings of Edmans et al. (2007) for an updated sample of stock market returns. The results of this preliminary analysis motivate a more in-depth intra-day analysis of investor sentiment. Accordingly, the analysis in Sub-section 3.4.2 examines the contemporaneous impact of football match outcomes on intra-day stock returns. The final portion of the analysis makes use of the trade and quote data described in Sub-section 3.3.4. The objective of this analysis is to directly test for abnormal buying, selling and quoting behaviour surrounding shocks to sentiment.

3.4.1. Daily Stock Return Analysis

The objective of this sub-section is to confirm the results of Edmans et al. (2007) for a more recent sample of daily stock market data. Accordingly, the empirical strategy of this sub-section follows from Edmans et al. (2007). The estimation is executed over two stages. The first stage involves regressing daily stock market returns across an array of control variables:

$$r_{m,d} = \alpha_m + \beta^1 r_{m,d-1} + \beta^2 r_{globe,d-1} + \beta^3 r_{globe,d} + \beta^4 r_{globe,d+1} + \beta^5 D_d + \beta^6 Q_{d-1} + \epsilon_{m,d} \tag{3.1}$$

where $r_{m,d}$ is the return on market m on day d, $r_{globe,d}$ is the daily US dollar return on Datastream's world market index on day d, $D_d = \{D_{1,d}, D_{2,d}, D_{3,d}, D_{4,d}\}$ are indicator variables for Monday through to Thursday and $Q_{d-1} = \{Q_{m,d-1}, Q_{m,d-2}, Q_{m,d-3}, Q_{m,d-4}, Q_{m,d-5}\}$ where $Q_{m,d-1}$ is an indicator variables that take the value of one if country m experienced a non-weekend public holiday on day $d-1$ and zero otherwise. Equation 3.1 is estimated simultaneously for all countries with country-level fixed effects and clustered errors.

The estimated residuals of Equation 3.1 are of key interest. Following Edmans et al. (2007), the estimated residuals are regressed on the sentiment variables derived from football match outcomes as such:

$$\begin{aligned}\epsilon_{m,d} =& \alpha + \beta^W \mathbb{1}\{W\}_{m,d} + \beta^L \mathbb{1}\{L\}_{m,d} + \beta^{GW} \mathbb{1}\{GW\}_{m,d} + \beta^{GL} \mathbb{1}\{GL\}_{m,d} \\ &+ \beta^{EW} \mathbb{1}\{EW\}_{m,d} + \beta^{EL} \mathbb{1}\{EL\}_{m,d} + u_{m,d}\end{aligned} \tag{3.2}$$

where $\mathbb{1}\{W\}_{m,d}$ is an indicator variable that takes the value of one if country m wins a World Cup football match on a day that makes d the first trading day after the match and zero otherwise and $\mathbb{1}\{L\}_{m,d}$ is an indicator variable analogously defined for losses. The indicator variable, $\mathbb{1}\{GW\}_{m,d}$, takes the value of one if country m wins a Group Stage World Cup football match on a day that makes d the first trading day after the match and zero otherwise, while $\mathbb{1}\{EW\}_{m,d}$ is analogously defined for Elimination Stage matches. The indicator variable, $\mathbb{1}\{GL\}_{m,d}$, takes the value of one if country m loses a Group Stage World Cup football match on a day that makes d the first trading day after the match and zero otherwise, while $\mathbb{1}\{EL\}_{m,d}$ is analogously defined for Elimination Stage matches.

Table 3.8
Abnormal Daily Stock Market Performance after International Football Matches

This table reports the estimation results for the following equation:

$$\epsilon_{m,d} = \alpha + \beta^{W}\mathbb{1}\{W\}_{m,d} + \beta^{L}\mathbb{1}\{L\}_{m,d} + \beta^{GW}\mathbb{1}\{GW\}_{m,d} + \beta^{GL}\mathbb{1}\{GL\}_{m,d} + \beta^{EW}\mathbb{1}\{EW\}_{m,d} + \beta^{EL}\mathbb{1}\{EL\}_{m,d} + u_{m,d} \tag{3.2}$$

where $\mathbb{1}\{W\}_{m,d}$ is an indicator variable that takes the value of one if country m wins a FIFA World Cup football match on a day that makes d the first trading day after the match and zero otherwise and $\mathbb{1}\{L\}_{m,d}$ is an indicator variable analogously defined for losses. The indicator variable, $\mathbb{1}\{GW\}_{m,d}$, takes the value of one if country m wins a Group Stage FIFA World Cup football match on a day that makes d the first trading day after the match and zero otherwise, while $\mathbb{1}\{EW\}_{m,d}$ is analogously defined for Elimination Stage matches. The indicator variable, $\mathbb{1}\{GL\}_{m,d}$, takes the value of one if country m loses a Group Stage FIFA World Cup football match on a day that makes d the first trading day after the match and zero otherwise, while $\mathbb{1}\{EL\}_{m,d}$ is analogously defined for Elimination Stage matches. The dependent variable, $\epsilon_{m,d}$, is defined by the following equation:

$$r_{m,d} = \alpha_m + \beta^1 r_{m,d-1} + \beta^2 r_{globe,d-1} + \beta^3 r_{globe,d} + \beta^4 r_{globe,d+1} + \beta^5 D_t + \beta^6 Q_{t-1} + \epsilon_{m,d} \tag{3.1}$$

where $r_{m,d}$ is the return on market m on day d, $r_{globe,d}$ is the daily U.S. dollar return on Datastream's world market index on day d, $D_t = \{D_{1,d}, D_{2,d}, D_{3,d}, D_{4,d}\}$ are indicator variables for Monday through to Thursday and $Q_{t-1} = \{Q_{m,d-1}, Q_{m,d-2}, Q_{m,d-3}, Q_{m,d-4}, Q_{m,d-5}\}$ where $Q_{m,d-1}$ is an indicator variables that take the value of one if country m experienced a non-weekend public holiday on day $d-1$ and zero otherwise. The t-statistics are reported in italics. The standard errors are clustered at the country level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively.

Year Range	No. of Games	β^{W}	No. of Games	β^{L}	No. of Games	β^{GW}	No. of Games	β^{GL}	No. of Games	β^{EW}	No. of Games	β^{EL}
1973-2014	320	0.014	248	-0.171**								
		0.21		*-2.18*								
					200	-0.047	147	-0.164*	120	0.116	101	-0.181
						-0.62		*-1.85*		*0.89*		*-1.59*
1973-2004	185	0.035	146	-0.330***								
		0.32		*-2.87*								
					108	0.008	84	-0.325**	77	0.073	62	-0.338**
						0.06		*-2.44*		*0.51*		*-1.98*
1973-1994	106	0.097	79	-0.345**								
		0.90		*-2.26*								
					57	0.081	43	-0.329**	49	0.117	36	-0.364*
						0.58		*-2.15*		*0.78*		*-1.73*
1995-2014	214	-0.026	169	-0.089								
		-0.34		*-1.03*								
					143	-0.096	104	-0.095	71	0.116	65	-0.079
						-0.95		*-0.99*		*0.76*		*-0.55*

Table 3.9

Abnormal Daily Stock Market Performance after International Football Matches for the Edmans et al. (2007) Sample of Countries

This table reports the estimation results for the following equation using the same sample of markets as Edmans et al. (2007):

$$\epsilon_{m,d} = \alpha + \beta^{W}\mathbb{1}\{W\}_{m,d} + \beta^{L}\mathbb{1}\{L\}_{m,d} + \beta^{GW}\mathbb{1}\{GW\}_{m,d} + \beta^{GL}\mathbb{1}\{GL\}_{m,d} + \beta^{EW}\mathbb{1}\{EW\}_{m,d} + \beta^{EL}\mathbb{1}\{EL\}_{m,d} + u_{m,d} \tag{3.2}$$

where $\mathbb{1}\{W\}_{m,d}$ is an indicator variable that takes the value of one if country m wins a FIFA World Cup football match on a day that makes d the first trading day after the match and zero otherwise and $\mathbb{1}\{L\}_{m,d}$ is an indicator variable analogously defined for losses. The indicator variable, $\mathbb{1}\{GW\}_{m,d}$, takes the value of one if country m wins a Group Stage FIFA World Cup football match on a day that makes d the first trading day after the match and zero otherwise, while $\mathbb{1}\{EW\}_{m,d}$ is analogously defined for Elimination Stage matches. The indicator variable, $\mathbb{1}\{GL\}_{m,d}$, takes the value of one if country m loses a Group Stage FIFA World Cup football match on a day that makes d the first trading day after the match and zero otherwise, while $\mathbb{1}\{EL\}_{m,d}$ is analogously defined for Elimination Stage matches. The dependent variable, $\epsilon_{m,d}$, is defined by the following equation:

$$r_{m,d} = \alpha_m + \beta^1 r_{m,d-1} + \beta^2 r_{globe,d-1} + \beta^3 r_{globe,d} + \beta^4 r_{globe,d+1} + \beta^5 D_t + \beta^6 Q_{t-1} + \epsilon_{m,d} \tag{3.1}$$

where $r_{m,d}$ is the return on market m on day d, $r_{globe,d}$ is the daily U.S. dollar return on Datastream's world market index on day d, $D_t = \{D_{1,d}, D_{2,d}, D_{3,d}, D_{4,d}\}$ are indicator variables for Monday through to Thursday and $Q_{t-1} = \{Q_{m,d-1}, Q_{m,d-2}, Q_{m,d-3}, Q_{m,d-4}, Q_{m,d-5}\}$ where $Q_{m,d-1}$ is an indicator variables that take the value of one if country m experienced a non-weekend public holiday on day $d-1$ and zero otherwise. The t-statistics are reported in italics. The standard errors are clustered at the country level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively.

Year Range	No. of Games	β^{W}	No. of Games	β^{L}	No. of Games	$\beta^{G,W}$	No. of Games	$\beta^{G,L}$	No. of Games	$\beta^{E,W}$	No. of Games	$\beta^{E,L}$
1973-2014	307	0.017	215	-0.239***								
		0.24		*-3.07*								
					189	-0.038	120	-0.268***	118	0.105	95	-0.201*
						-0.50		*-3.18*		*0.79*		*-1.67*
1973-2004	181	0.043	129	-0.378***								
		0.38		*-3.17*								
					105	0.027	69	-0.404***	76	0.066	60	-0.348**
						0.21		*-2.95*		*0.45*		*-1.97*
1973-1994	105	0.105	71	-0.378**								
		0.97		*-2.31*								
					56	0.097	36	-0.389**	49	0.115	35	-0.367*
						0.69		*-2.41*		*0.76*		*-1.70*
1995-2014	202	-0.029	144	-0.169*								
		-0.37		*-1.90*								
					133	-0.094	84	-0.215**	69	0.096	60	-0.104
						-0.90		*-2.17*		*0.62*		*-0.67*

Table 3.8 presents the estimation results of Equation 3.2. The results presented in Table 3.8 indicate that wins do not have a significant impact on daily abnormal stock returns. This is consistent with Edmans et al. (2007). Nonetheless, the results for the entire sample period, 1973 to 2014, indicate that there is a strong loss-effect that can be attributed to negative sentiment. Following a loss in a World Cup football match, markets exhibit an abnormal stock return of -17.1 basis points. This result is significant at the 95% confidence level. Further, after a loss in a World Cup Group Stage match, markets experience a -16.4 basis point loss effect. The analogous impact of an Elimination Stage loss is marginally insignificant.

Table 3.8 also presents the results of Equation 3.2 for a variety of time periods. The 1973 to 2004 range coincides with the original sample period of Edmans et al. (2007). Table 3.8 demonstrates that the loss-effect is considerably stronger for the Edmans et al. (2007) sample period. For the 1973 to 2004 time period, losses are associated with a -33.0 basis point abnormal stock return. This result is significant at the 99% confidence level. Further, the loss-effect is statistically significant at the 95% confidence level for both Group Stage and Elimination Stage matches when the analysis is restricted to 1973 to 2004.

Table 3.8 also considers the estimation results of Equation 3.2 for the 1973 to 1994 and 1995 to 2014 sample periods. The results presented in Table 3.8 indicate that the loss-effect is strongest for the first 20 years of data and less so for the most recent 20 years. The estimated β^{L}, β^{GL} and β^{EL} coefficients are all negative and statistically significant from zero for the 1973 to 1994 time period. In contrast, the estimated β^{L}, β^{GL} and β^{EL} coefficients are insignificant from zero for the 1995 to 2014 time period. Thus, the results suggest that the behavioural anomaly has weakened over time.

Table 3.9 gives the Equation 3.2 estimation results for the countries present in the Edmans et al. (2007) study. Table 3.9 demonstrates that the loss-effect results are generally stronger for the Edmans et al. (2007) set of countries. In particular, the estimated β^{L} coefficient for the most recent time period, 1995 to 2014, is -0.169 and significantly different from zero at the 90% confidence level. For the entire set of countries, the analogous figure in Table 3.8 is -0.095 and statistically insignificant from zero. Thus, for the Edmans et al. (2007) subset of countries, the loss effect is prevalent within the most recent 20 years of stock market data.

3.4.2. Five-minute Intra-day Stock Return Analysis

On occasion, a country's national stock exchange may be open for trading while the country's national football team is participating in a World Cup football match. This motivates an investigation of the impacts of intra-day shocks to investor sentiment. This is achieved by using the five-minute intra-day stock market returns dataset described in Sub-section 3.3.3 and a methodology similar in design to Edmans et al. (2007). The intra-day analysis is conducted at the five-minute frequency to avoid timing issues regarding football matches that do not run exactly to schedule.

Similar to the daily stock market returns analysis, the estimation procedure of the intra-day stock return analysis is comprised of two stages. The first stage utilises the methodology of Gallant et al. (1992) to control for seasonal trends in the data. This seasonality adjustment maintains the mean and variance of the data, after removing all variation that is explained by the seasonal variables. The seasonal variables consist of: month-of-the-year indicator variables; day-of-the-week indicator variables; and, five-minute time-of-the-day indicator variables. I individually and independently implement the Gallant et al. (1992) procedure for each country.

The second stage of the empirical analysis involves estimating the following equation:

$$r_{m,t} = \alpha_m + \beta^{globe} r_{globe,t} + \beta^{HW} \mathbb{1}\{HW\}_{m,t} + \beta^{HL} \mathbb{1}\{HL\}_{m,t} + \beta^{W} \mathbb{1}\{W\}_{m,t} + \beta^{L} \mathbb{1}\{L\}_{m,t} + \epsilon_{m,t} \tag{3.3}$$

where $r_{m,t}$ is the seasonally detrended five-minute intra-day return of market m at time t and $r_{globe,t}$ is the seasonally detrended five-minute return on the S&P Global Broad Market Index at time t.[28] The indicator variable, $\mathbb{1}\{HW\}_{m,t}$, takes the value of one during half-time of a football match in which the national football team associated with market m is participating, conditional on that team winning at half-time. The indicator variable, $\mathbb{1}\{HL\}_{m,t}$, takes the value of one during half-time of a football match in which the national football team associated with market m is participating, conditional on that team losing at half-time. The indicator variable, $\mathbb{1}\{W\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m winning the match. The indicator variable, $\mathbb{1}\{L\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m losing the match. According to the investor sentiment literature, we should expect the β^{HW} and β^{W} coefficients to be positive, following a positive shock to sentiment. Conversely, we should expect the β^{HL} and β^{L} coefficients to be negative.

Table 3.10 presents the estimation results of Equation 3.3. The signs of the estimated β^{HW} and β^{HL} coefficients are in-line with the sentiment hypothesis but insignificantly different from zero. The estimated β^{W} coefficients provide evidence of a sentiment effect. The estimated β^{W} coefficients range from 0.5 to 0.6 basis points and are significantly different from zero. The magnitude of this coefficient is economically significant The estimated β^{W} coefficients correspond to an abnormal 0.5-0.6 basis point return every five minutes between full-time and the close of trading. The average time between a winning full-time outcome and the close of trade is 143.13 minutes. Thus, the estimated β^{W} coefficients equate to an abnormal cumulative stock return from full-time to the close of trading of between 14.3 and 17.2 basis points. Perhaps surprisingly, the estimated β^{L} coefficients presented in Table 3.10 are not significantly different from zero. Thus, in contrast to the Equation 3.2 estimation results, the Equation 3.3 estimation results provide evidence of a win-effect and *do not* provide evidence of a loss-effect. These results can be reconciled if one considers the dynamics of the contrasting sentiment effects. Cai et al. (2018) argue

[28]The $r_{globe,t}$ variable is available from January 2000.

Table 3.10
Abnormal Intra-Day Stock Market Performance during International Football Matches

This table reports the estimation results for the following regression:

$$r_{m,t} = \alpha_m + \beta^{globe} r_{globe,t} + \beta^{HW} \mathbb{1}\{HW\}_{m,t} + \beta^{HL} \mathbb{1}\{HL\}_{m,t} + \beta^{W} \mathbb{1}\{W\}_{m,t} + \beta^{L} \mathbb{1}\{L\}_{m,t} + \epsilon_{m,t} \quad (3.3)$$

where $r_{m,t}$ is the seasonally detrended five-minute intra-day return of market m at time t and $r_{globe,t}$ is the seasonally detrended five-minute return on the S&P Global Broad Market Index at time t. The indicator variables, $\mathbb{1}\{HW\}_{m,t}$, takes the value of one during half-time of a football match in which the national football team associated with market m is participating, conditional on that team winning at half-time. The indicator variable, $\mathbb{1}\{HL\}_{m,t}$, takes the value of one during half-time of a football match in which the national football team associated with market m is participating, conditional on that team losing at half-time. The indicator variable, $\mathbb{1}\{W\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m winning the match. The indicator variable, $\mathbb{1}\{L\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m losing the match. The t-statistics are reported in italics. The standard errors are clustered at the country level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively. R^2_W is a weighted coefficient of determination that only utilises observations for which at least one match outcome indicator variable is nonzero.

	(1)	(2)
β^{globe}		0.056*
		1.72
β^{HW}	0.032	0.037
	1.62	*1.56*
β^{HL}	-0.007	-0.008
	-0.72	*-0.69*
β^{W}	0.005*	0.006*
	1.93	*1.81*
β^{L}	0.002	0.001
	0.52	*0.15*
Observations	7144776	6360781
R^2_W (%)	0.44	0.55

that the overnight loss-effect is partially attributed to the physiological impacts of sleeplessness that may result if investors disrupt their regular sleeping patterns to monitor football matches. Football matches that occur during regular trading hours should not be expected to disrupt sleeping patterns. Therefore the predominant win-effect resultant from intra-day match outcomes does not directly contradict the results presented in Section 3.4.1 and Edmans et al. (2007).

Figure 3.1 provides a graphical representation of the intra-day returns following full-time. Figure 3.1 distinguishes between winning and losing full-time match outcomes. Panels A and B plot the mean seasonally detrended five-minute intra-day stock returns following a winning and losing full-time match outcome, respectively. At each time period, the returns are mostly insignificantly different from zero. This is likely due to the lack of observations at each individual time period. Accordingly, to enable a more straightforward interpretation, I construct cumulative returns from panels A and B. Panel C of Figure 3.1 plots the cumulative returns derived from the mean returns presented in Panel A. Panel D plots the cumulative returns from the mean returns presented in

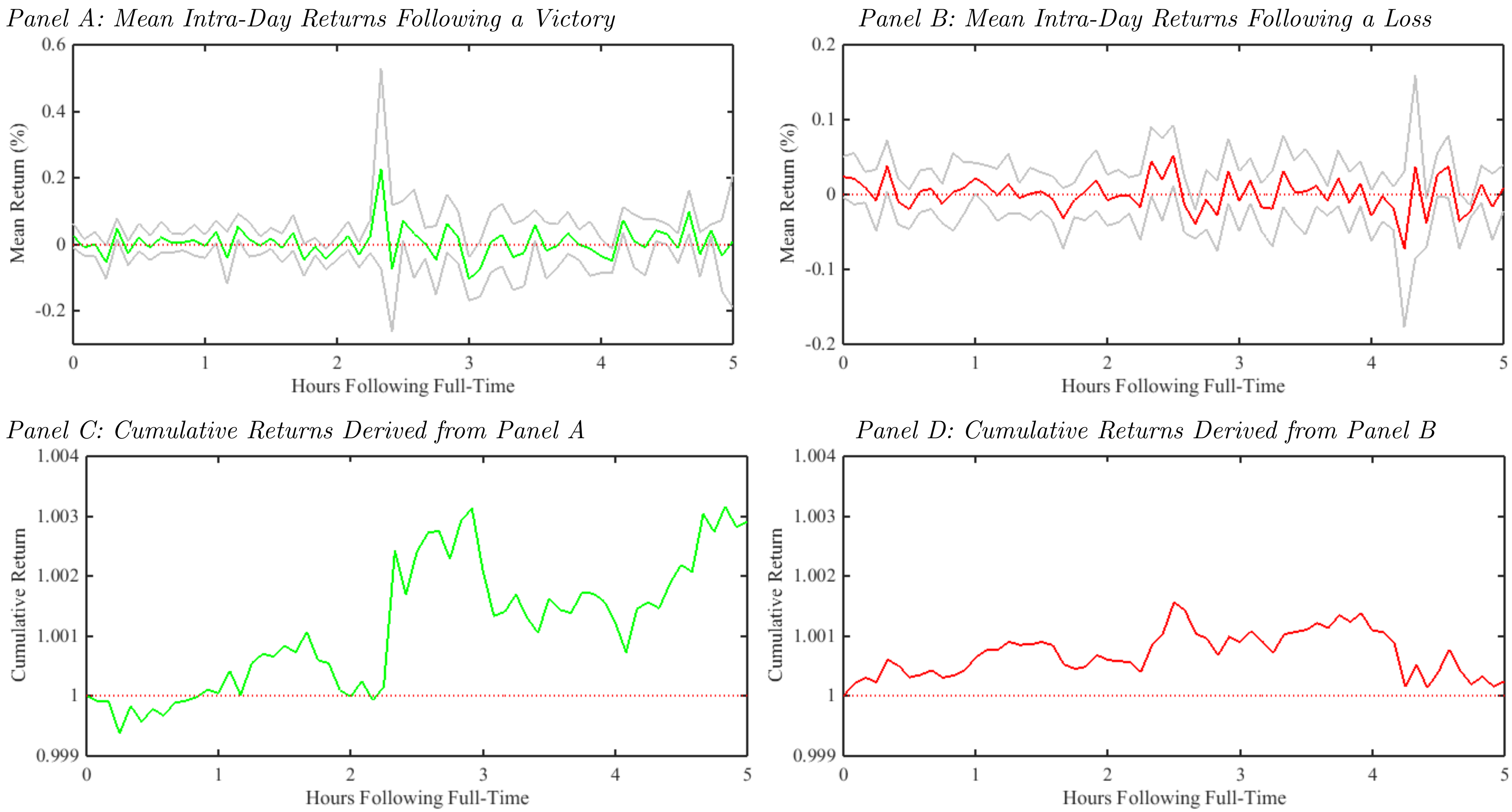


Fig. 3.1. Intra-day Return following Full-time. Panel A plots the mean seasonally detrended five-minute intra-day stock returns following a winning full-time match outcome, as well as the associated 90% confidence interval. The market variables are detrended using the methodology of Gallant et al. (1992) for month-of-the-year, day-of-the-week and five-minute time-of-the-day effects. Panel B is the equivalent of Panel A for losing match outcomes. Panel C plots a cumulative returns derived from the mean returns presented in Panel A. Panel D is the equivalent of Panel C for losses and Panel B.

Panel B. The returns presented in panels C and D should be interpreted as "artificial" cumulative returns as they are derived from mean return series rather than realised return series.[29] Nonetheless, Panel C is representative of the positive abnormal returns that occur following a full-time winning match outcome. Panel C depicts a cumulative return of approximately 30 basis points after five hours from full-time. In contrast and in-line with the statistical results presented in Table 3.10, Panel D does not allude to any meaningful cumulative returns following losing full-time match outcomes.

3.4.2.1 Group Stage and Elimination Stage Matches

Similar to the daily stock market returns analysis, it is also possible to consider the separate impacts of Group Stage match outcomes and Elimination Stage match outcomes at the intra-day level. Accordingly, I consider the following equation:

$$\begin{aligned} r_{m,t} =& \alpha_m + \beta^{globe} r_{globe,t} + \beta^{HW} \mathbb{1}\{HW\}_{m,t} + \beta^{HL} \mathbb{1}\{HL\}_{m,t} + \beta^{GHW} \mathbb{1}\{GHW\}_{m,t} \\ &+ \beta^{GHL} \mathbb{1}\{GHL\}_{m,t} + \beta^{W} \mathbb{1}\{W\}_{m,t} + \beta^{L} \mathbb{1}\{L\}_{m,t} + \beta^{GW} \mathbb{1}\{GW\}_{m,t} + \beta^{GL} \mathbb{1}\{GL\}_{m,t} \\ &+ \beta^{EW} \mathbb{1}\{EW\}_{m,t} + \beta^{EL} \mathbb{1}\{EL\}_{m,t} + \epsilon_{m,t} \end{aligned} \tag{3.4}$$

where the indicator variables, $\mathbb{1}\{GHW\}_{m,t}$ and $\mathbb{1}\{EHW\}_{m,t}$, are the equivalents of $\mathbb{1}\{HW\}_{m,t}$ with respect to Group Stage matches and Elimination Stage matches, respectively. The indicator variables, $\mathbb{1}\{GHL\}_{m,t}$ and $\mathbb{1}\{EHL\}_{m,t}$, are the equivalents of $\mathbb{1}\{HL\}_{m,t}$ with respect to Group Stage matches and Elimination Stage matches, respectively. The indicator variables, $\mathbb{1}\{GW\}_{m,t}$ and $\mathbb{1}\{EW\}_{m,t}$, are the equivalents of $\mathbb{1}\{W\}_{m,t}$ with respect to Group Stage matches and Elimination Stage matches, respectively. The indicator variables, $\mathbb{1}\{GL\}_{m,t}$ and $\mathbb{1}\{EL\}_{m,t}$, are the equivalents of $\mathbb{1}\{L\}_{m,t}$ with respect to Group Stage matches and Elimination Stage matches, respectively.

The estimated results for Equation 3.4 are presented in Table 3.11. All estimated coefficients presented in Table 3.11 pertaining to half-time match outcomes are of the sign that is predicted by the sentiment literature. That is, the estimated β^{GHW} and β^{EHW} coefficients are positive; while, the estimated β^{GHL} and β^{EHL} coefficients are negative. The estimated β^{EHL} coefficient for the third specification presented in Table 3.11 is -0.130 with a remarkable t-statistic of -1162.77. This is because there is only one observation of an Elimination Stage half-time win after the sample period of intra-day stock returns is reduced to accommodate the data availability of the $r_{globe,t}$ variable. The robust t-statistic of this coefficient, given by a counter-factual analysis, is -2.50. This analysis is discussed in Appendix 3.6 and presented in Table 3.20.

The estimated coefficients of Equation 3.4 pertaining to full-time match outcomes provide mixed evidence of the sentiment hypothesis. That is, the estimated β^{GW}, β^{GL} and β^{EW} coefficients

[29] It is also possible to plot mean cumulative returns following full-time. Unfortunately, this plot suffers from a small sample problem. This is because the time interval between full-time and the close of trade is variable. Accordingly, the mean cumulative returns series have jumps corresponding to points in time where the sample size reduces. This inhibits a reasonable interpretation of the series.

Table 3.11
Abnormal Intra-Day Stock Market Performance during World Cup Group Stage and Elimination Stage Football Matches

This table reports the estimation results for the following regression:

$$\begin{aligned} r_{m,t} = \alpha_m + \beta^{globe} r_{globe,t} + \beta^{HW} \mathbb{1}\{HW\}_{m,t} + \beta^{HL} \mathbb{1}\{HL\}_{m,t} + \beta^{GHW} \mathbb{1}\{GHW\}_{m,t} + \beta^{GHL} \mathbb{1}\{GHL\}_{m,t} \\ + \beta^{W} \mathbb{1}\{W\}_{m,t} + \beta^{L} \mathbb{1}\{L\}_{m,t} + \beta^{GW} \mathbb{1}\{GW\}_{m,t} + \beta^{GL} \mathbb{1}\{GL\}_{m,t} + \beta^{EW} \mathbb{1}\{EW\}_{m,t} + \beta^{EL} \mathbb{1}\{EL\}_{m,t} + \epsilon_{m,t} \end{aligned} \quad (3.4)$$

where $r_{m,t}$ is the seasonally detrended five-minute intra-day return of market m at time t and $r_{globe,t}$ is the seasonally detrended five-minute return on the S&P Global Broad Market Index at time t. The indicator variable, $\mathbb{1}\{HW\}_{m,t}$, takes the value of one during half-time of a football match in which the national football team associated with market m is participating, conditional on that team winning at half-time. The indicator variables, $\mathbb{1}\{GHW\}_{m,t}$ and $\mathbb{1}\{EHW\}_{m,t}$, are the equivalents of $\mathbb{1}\{HW\}_{m,t}$ with respect to Group Stage matches and Elimination Stage matches, respectively. The indicator variables, $\mathbb{1}\{HL\}_{m,t}$, takes the value of one during half-time of a football match in which the national football team associated with market m is participating, conditional on that team losing at half-time. The indicator variables, $\mathbb{1}\{GHL\}_{m,t}$ and $\mathbb{1}\{EHL\}_{m,t}$, are the equivalents of $\mathbb{1}\{HL\}_{m,t}$ with respect to Group Stage matches and Elimination Stage matches, respectively. The indicator variable, $\mathbb{1}\{W\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m winning the match. The indicator variables, $\mathbb{1}\{GW\}_{m,t}$ and $\mathbb{1}\{EW\}_{m,t}$, are the equivalents of $\mathbb{1}\{W\}_{m,t}$ with respect to Group Stage matches and Elimination Stage matches, respectively. The indicator variable, $\mathbb{1}\{L\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m losing the match. The indicator variables, $\mathbb{1}\{GL\}_{m,t}$ and $\mathbb{1}\{EL\}_{m,t}$, are the equivalents of $\mathbb{1}\{L\}_{m,t}$ with respect to Group Stage matches and Elimination Stage matches, respectively. The t-statistics are reported in italics. The standard errors are clustered at the country level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively. R^2_W is a weighted coefficient of determination that only utilises observations for which at least one match outcome indicator variable is nonzero.

	(1)	(2)	(3)	(4)
β^{globe}			0.056*	0.056*
			1.72	*1.72*
β^{HW}		0.032		0.037
		1.62		*1.56*
β^{HL}		-0.007		-0.008
		-0.72		*-0.69*
β^{GHW}	0.021		0.020	
	1.47		*1.13*	
β^{GHL}	-0.001		-0.001	
	-0.14		*-0.11*	
β^{EHW}	0.071		0.089	
	1.15		*1.49*	
β^{EHL}	-0.064		-0.130***	
	-1.35		*-1162.77*	
β^{W}	0.005*		0.006*	
	1.93		*1.81*	
β^{L}	0.002		0.001	
	0.52		*0.15*	
β^{GW}		0.002		0.003
		0.43		*0.56*
β^{GL}		-0.002		-0.003
		-0.43		*-0.61*
β^{EW}		0.015**		0.018*
		2.17		*1.74*
β^{EL}		0.011**		0.012*
		2.24		*1.71*
Observations	7144776	7144776	6360781	6360781
R^2_W (%)	0.65	0.63	0.97	0.79

conform to the sentiment predictions; however, the estimated β^{EL} coefficients oppose the sentiment predictions. The estimated β^{GW} and β^{EW} coefficients are positive, indicating positive abnormal stock returns following a victory. The estimate β^{EW} coefficients are statistically significant from zero at the 90% confidence level. The signs of the estimated β^{GL} coefficients are negative but insignificantly different from zero. Contrary to the sentiment predictions, the estimated β^{EL} coefficients are positive and statistically significant. This surprising result could be due to the small number of Elimination Stage loss observations in the sample. As discussed in Sub-section 3.3.3, Table 3.5 demonstrates that there are only six full-time Elimination Stage loss observations in the sample. Further analysis presented in Appendix 3.6.2 demonstrates that the estimated β^{EL} coefficients are not robust to a rigorous placebo test presented in Table 3.20. The robust t-statistics for the estimated β^{EL} coefficients in Table 3.20 are 1.35 and 1.10 for specifications (2) and (4) of Equation 3.4, respectively. This suggests that the estimated β^{EL} coefficients are driven by outlier events.

3.4.2.2 Unexpected Match Outcomes

One could argue that if investors form conditional expectations regarding future uncertain events, *unexpected* match outcomes might be a more precise tool for identifying shocks to investor sentiment. For example, an unexpected winning full-time match outcome may result in a greater proliferation of positive sentiment than an expected winning full-time match outcome. Accordingly, I utilise the Elo (1978) rating system to stratify the post-1994 match outcomes in the sample into expected and unexpected match outcomes. The Elo (1978) rating system is used to rank players in repeated zero-sum games. The Elo (1978) rating system gained prominence in the world of competitive chess but can equally be applied to competitive football. As international football Elo (1978) ratings are time-varying, I extract Elo (1978) ratings for each country in the sample, prior to each iteration of the World Cup.[30] Following Edmans et al. (2007), I then use a 125 point Elo (1978) rating differential to identify unexpected results. A match outcome is categorised as an unexpected result if the losing team has an Elo (1978) rating that is more than 125 points greater than that of the winning team. Table 3.12 presents the post-1994 unexpected match outcomes that are identified by this methodology. Within the sample, there are 11 unexpected half-time match outcomes and nine unexpected full-time match outcomes.

To test whether unexpected match outcomes have a significant marginal impact on intra-day stock returns, I estimate the following equation:

$$\begin{aligned} r_{m,t} =& \alpha_m + \beta^{globe} r_{globe,t} + \beta^{HW}\mathbb{1}\{HW\}_{m,t} + \beta^{HL}\mathbb{1}\{HL\}_{m,t} + \beta^{UHW}\mathbb{1}\{UHW\}_{m,t} \\ &+ \beta^{UHL}\mathbb{1}\{UHL\}_{m,t} + \beta^{W}\mathbb{1}\{W\}_{m,t} + \beta^{L}\mathbb{1}\{L\}_{m,t} + \beta^{UW}\mathbb{1}\{UW\}_{m,t} + \beta^{UL}\mathbb{1}\{UL\}_{m,t} + \epsilon_{m,t} \end{aligned} \tag{3.5}$$

[30] Elo (1978) ratings are extracted from www.eloratings.net.

Table 3.12
Post-1994 Unexpected Match Outcomes

This table summarises the post-1994 unexpected match outcomes within the football match sample. Following Edmans et al. (2007), I utilise Elo (1978) ratings to identify unexpected results. To categorise unexpected match outcomes, I collect every country's Elo (1978) rating at the start of each FIFA World Cup from www.eloratings.net. A match outcome is categorised as an unexpected result if the losing team has an Elo (1978) rating that is more than 125 points greater than that of the winning team.

Unexpected Half-time Match Outcomes				
Date	Country within Sample	Goals	Opposition Country	Goals
04-Jul-1998	Germany	0	Croatia	1
31-May-2002	France	0	Senegal	1
05-Jun-2002	Portugal	1	United States of America	3
05-Jun-2002	United States of America	3	Portugal	1
07-Jun-2002	Spain	0	Paraguay	1
11-Jun-2002	France	0	Denmark	1
13-Jun-2002	Mexico	1	Italy	0
19-Jun-2006	Spain	0	Tunisia	1
22-Jun-2010	France	0	South Africa	2
22-Jun-2010	South Africa	2	France	0
18-Jun-2014	Chile	2	Spain	0

Unexpected Full-time Match Outcomes				
Date	Country within Sample	Goals	Opposition Country	Goals
31-May-2002	France	0	Senegal	1
05-Jun-2002	Portugal	2	United States of America	3
05-Jun-2002	United States of America	3	Portugal	2
11-Jun-2002	France	0	Denmark	2
14-Jun-2002	Portugal	0	Republic of Korea	1
18-Jun-2002	Italy	1	Republic of Korea	2
22-Jun-2006	United States of America	1	Ghana	2
22-Jun-2010	France	1	South Africa	2
22-Jun-2010	South Africa	2	France	1

where the indicator variable, $\mathbb{1}\{UHW\}_{m,t}$, is the equivalent of $\mathbb{1}\{HW\}_{m,t}$ with respect to unexpected winning half-time match outcomes and the indicator variable, $\mathbb{1}\{UHL\}_{m,t}$, is the equivalent of $\mathbb{1}\{HL\}_{m,t}$ with respect to unexpected losing half-time match outcomes. The indicator variable, $\mathbb{1}\{UW\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m unexpectedly winning the match. The indicator variable, $\mathbb{1}\{UL\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m unexpectedly losing the match.

The estimation results of Equation 3.5 are presented in Table 3.13. The results demonstrate that while unexpected half-time and full-time loss match outcomes have an insignificant marginal impact on intra-day stock returns, unexpected full-time winning match outcomes have a positive and statistically significant marginal impact. The estimated β^{UW} coefficients presented in Table 3.13 are statistically significant at the 99% confidence level and range between 0.007 and 0.010. Thus, an unexpected victory has a 0.7 to 1.0 basis point marginal impact on five-minute intra-day stock

Table 3.13
Abnormal Intra-Day Stock Market Performance during Unexpected World Cup Match Outcomes

This table reports the estimation results for the following regression:

$$\begin{aligned} r_{m,t} =& \alpha_m + \beta^{globe} r_{globe,t} + \beta^{HW}\mathbb{1}\{HW\}_{m,t} + \beta^{HL}\mathbb{1}\{HL\}_{m,t} + \beta^{UHW}\mathbb{1}\{UHW\}_{m,t} + \beta^{UHL}\mathbb{1}\{UHL\}_{m,t} \\ &+ \beta^{W}\mathbb{1}\{W\}_{m,t} + \beta^{L}\mathbb{1}\{L\}_{m,t} + \beta^{UW}\mathbb{1}\{UW\}_{m,t} + \beta^{UL}\mathbb{1}\{UL\}_{m,t} + \epsilon_{m,t} \end{aligned} \tag{3.5}$$

where $r_{m,t}$ is the seasonally detrended five-minute intra-day return of market m at time t and $r_{globe,t}$ is the seasonally detrended five-minute return on the S&P Global Broad Market Index at time t. The indicator variable, $\mathbb{1}\{HW\}_{m,t}$, takes the value of 1 during half-time of a football match in which the national football team associated with market m is participating, conditional on that team winning at half-time. The indicator variables, $\mathbb{1}\{HL\}_{m,t}$, takes the value of one during half-time of a football match in which the national football team associated with market m is participating, conditional on that team losing at half-time. The indicator variable, $\mathbb{1}\{UHW\}_{m,t}$, is the equivalent of $\mathbb{1}\{HW\}_{m,t}$ with respect to unexpected winning half-time match outcomes and the indicator variable, $\mathbb{1}\{UHL\}_{m,t}$, is the equivalent of $\mathbb{1}\{HL\}_{m,t}$ with respect to unexpected losing half-time match outcomes. The indicator variable, $\mathbb{1}\{W\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m winning the match. The indicator variable, $\mathbb{1}\{L\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m losing the match. The indicator variable, $\mathbb{1}\{UW\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m unexpectedly winning the match. The indicator variable, $\mathbb{1}\{UL\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m unexpectedly losing the match. A win (loss) is characterised as unexpected if the winning (losing) team had an ELO score that was at least 125 lower (higher) than the ELO score of the losing (winning) team, at the commencement of the respective World Cup. The t-statistics are reported in italics. The standard errors are clustered at the country level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively. R^2_W is a weighted coefficient of determination that only utilises observations for which at least one match outcome indicator variable is nonzero.

	(1)	(2)	(3)	(4)
β^{globe}			0.056*	0.056*
			1.72	*1.72*
β^{HW}	0.032	0.032	0.036	0.037
	1.50	*1.62*	*1.47*	*1.56*
β^{HL}	-0.010	-0.007	-0.014	-0.008
	-0.75	*-0.72*	*-0.75*	*-0.69*
β^{UHW}	0.004		0.017	
	0.18		*0.69*	
β^{UHL}	0.011		0.015	
	0.69		*0.73*	
β^{W}	0.005*	0.005*	0.006*	0.006*
	1.93	*1.93*	*1.81*	*1.81*
β^{L}	0.002	0.001	0.001	-0.001
	0.52	*0.21*	*0.15*	*-0.23*
β^{UW}		0.010***		0.007***
		11.50		*4.72*
β^{UL}		0.004		0.006
		0.74		*1.09*
Observations	7144776	7144776	6360781	6360781
R^2_W (%)	0.26	0.28	0.37	0.38

returns from full-time to the close of trading on match days. Given that the average time between an unexpected winning full-time match outcome and market close is 182.22 minutes, the 0.007 and 0.010 estimated coefficients equate to a cumulative return of between 25.6 and 36.5 basis points.

3.4.2.3 Traditional Football Rivalries

One could argue that external historical factors influence the impact of football match outcomes on investor sentiment. For example, within football, a "traditional" rivalry is a rivalry that is perceived to be of great importance. Traditional rivalries are often attributed to past historical events or the geographical proximity of the competing countries. One example is the traditional football rivalry between Argentina and England. This rivalry is often attributed to the Falklands War. Accordingly, it could be hypothesised that a victory over a major rival induces greater positive sentiment than a victory over a non-traditional rival. This could mean that victories over traditional rivals have a significant marginal impact on abnormal stock returns. To test this notion, I construct an *a priori* set of traditional rivals for each country in the sample. The rivalries are presented in Table 3.14.

From the list of traditional rivalries presented in Table 3.14, I find that there are a number of winning full-time match outcomes over traditional rivals that occur during trading hours. There are no losses to traditional rivals that occur during trading hours. Thus, I augment Equation 3.3 to include an additional indicator variable, $\mathbb{1}\{RW\}_{m,t}$:

$$\begin{aligned} r_{m,t} =& \alpha_m + \beta^{globe} r_{globe,t} + \beta^{HW}\mathbb{1}\{HW\}_{m,t} + \beta^{HL}\mathbb{1}\{HL\}_{m,t} \\ &+ \beta^{W}\mathbb{1}\{W\}_{m,t} + \beta^{L}\mathbb{1}\{L\}_{m,t} + \beta^{RW}\mathbb{1}\{RW\}_{m,t} + \epsilon_{m,t} \end{aligned} \tag{3.6}$$

The indicator variable, $\mathbb{1}\{RW\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m winning the match over a traditional rival.

Table 3.15 presents the estimation results for Equation 3.6. The results reveal that winning over a traditional rival has a very significant marginal impact on abnormal stock returns for the rest of the trading day. Winning over a traditional rival has a 0.9 to 1.0 basis point marginal impact on stock returns every five-minutes for the period between full-time and the close of trading on match day. For both specifications of Equation 3.6 presented in Table 3.15, the estimated β^{RW} coefficients are significantly different from zero at the 95% confidence level. Given that the average time between a winning full-time match outcome over a traditional rival and the close of trade is 130.00 minutes, the estimated β^{RW} coefficients translate to an abnormal cumulative return of between 23.4 and 26.03 basis points. Thus, the results presented in Table 3.15 provide additional evidence in favour of the sentiment hypotheses.

Table 3.14
Traditional Football Rivalries

This table gives the traditional rivals of the countries in the intra-day analysis.

Country	Traditional Rivals			
Argentina	Brazil	Chile	Uruguay	
Belgium	Netherlands			
Brazil	Argentina			
Chile	Argentina	Peru		
Denmark	Norway	Sweden		
England	Argentina	Germany	Republic of Ireland	Scotland
France				
Germany	England	Netherlands		
Republic of Ireland	England			
Italy				
Mexico	United States of America			
Netherlands	Belgium	Germany		
Poland	Russia			
Portugal	Spain			
South Africa				
Spain	Portugal			
Switzerland				
Turkey	Armenia	Croatia	Greece	
United States of America				

3.4.3. Trade and Quote Analysis

The preceding analyses have demonstrated the correlation between World Cup football match outcomes and abnormal stock market returns. The behavioural finance literature argues that this correlation is caused by shifting levels of investor optimism, which in turn influences investor buying and selling decisions. The objective of this sub-section is to directly test this argument by observing buying, selling and quoting activities.

3.4.3.1 Order Imbalance Analysis

To observe buying and selling behaviour, I apply the Lee and Ready (1991) algorithm to the trade and quote dataset. This enables me to categorise trades as buyer-initiated or seller-initiated trades. Following, I calculate the three order imbalance measures of Chordia et al. (2002). The first measure is denoted by $OIBNUM_{m,t,w}$ for market m, five-minute intra-day time period t and World Cup sub-sample w. The $OIBNUM_{m,t,w}$ measure, is the number of buyer-initiated trades less the number of seller-initiated trades on market m at time t in World Cup sub-sample w:

$$OIBNUM_{m,t,w} = BUYS_{m,t,w} - SELLS_{m,t,w}. \tag{3.7}$$

The other order imbalance measures are $OIBSH_{m,t,w}$, the number of buyer-initiated shares purchased less the number the seller-initiated shares sold and $OIBDOL_{m,t,w}$, the buyer-initiated dollars

Table 3.15
Abnormal Intra-Day Stock Market Performance during International Football Matches between Traditional Rivals

This table reports the estimation results for the following regression:

$$\begin{aligned} r_{m,t} =& \alpha_m + \beta^{globe} r_{globe,t} + \beta^{HW} \mathbb{1}\{HW\}_{m,t} + \beta^{HL} \mathbb{1}\{HL\}_{m,t} \\ &+ \beta^{W} \mathbb{1}\{W\}_{m,t} + \beta^{L} \mathbb{1}\{L\}_{m,t} + \beta^{RW} \mathbb{1}\{RW\}_{m,t} + \epsilon_{m,t} \end{aligned} \tag{3.6}$$

where $r_{m,t}$ is the seasonally detrended five-minute intra-day return of market m at time t and $r_{globe,t}$ is the seasonally detrended five-minute return on the S&P Global Broad Market Index at time t. The indicator variable, $\mathbb{1}\{HW\}_{m,t}$, takes the value of one during half-time of a football match in which the national football team associated with market m is participating, conditional on that team winning at half-time. The indicator variables, $\mathbb{1}\{HL\}_{m,t}$, takes the value of one during half-time of a football match in which the national football team associated with market m is participating, conditional on that team losing at half-time. The indicator variable, $\mathbb{1}\{W\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m winning the match. The indicator variable, $\mathbb{1}\{L\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m losing the match. The indicator variable, $\mathbb{1}\{RW\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m winning the match over a traditional rival. The t-statistics are reported in italics. The standard errors are clustered at the country level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively. R^2_W is a weighted coefficient of determination that only utilises observations for which at least one match outcome indicator variable is nonzero.

	(1)	(2)
β^{globe}		0.056^{*}
		1.72
β^{HW}	0.032	0.037
	1.62	*1.56*
β^{HL}	-0.007	-0.008
	-0.72	*-0.69*
β^{W}	0.005^{*}	0.006^{*}
	1.75	*1.70*
β^{L}	0.002	0.001
	0.52	*0.15*
β^{RW}	0.010^{***}	0.009^{**}
	3.83	*2.49*
Observations	7144776	6360781
R^2_W (%)	0.28	0.38

(or local currency) paid less the seller-initiated dollars (or local currency) received, both for market m, time t and World Cup sub-sample w.

I make two adjustments to the $OIBNUM_{m,t,w}$, $OIBSH_{m,t,w}$ and $OIBDOL_{m,t,w}$ data to allow for robust economic and statistical inference. First, I standardise each variable to have a mean of zero and standard deviation of one. This is because the magnitude of each variable is not consistent over the cross-section. This is due to the fact that each market index is comprised of a unique number of constituents and is often denominated in a unique currency. Second, I use the Gallant et al. (1992) detrending process to remove all variation in the data that can be explained by seasonal and intra-day variables. Similar to the previous sub-section, the seasonal variables include month-of-the-year indicator variables; day-of-the-week indicator variables; and, five-minute time-of-the-day indicator variables. For simplicity, for the remainder of the paper let $OIBNUM_{m,t,w}$, $OIBSH_{m,t,w}$ and $OIBDOL_{m,t,w}$ refer to the adjusted variables.

To determine whether investor sentiment influences buying and selling behaviour, I estimate the following equation:

$$\begin{aligned} y_{m,t,w} =& \alpha_{m,w} + \beta^{globe} r_{globe,t,w} + \beta^{HW} \mathbb{1}\{HW\}_{m,t,w} + \beta^{HL} \mathbb{1}\{HL\}_{m,t,w} \\ &+ \beta^{W} \mathbb{1}\{W\}_{m,t,w} + \beta^{L} \mathbb{1}\{L\}_{m,t,w} + \epsilon_{m,t,w} \end{aligned} \tag{3.8}$$

where $y_{m,t,w}$ is the seasonally detrended and standardised dependent variable of market m at five-minute intra-day time period t and World Cup sub-sample w. Equation 3.8 is simultaneously estimated for all countries and World Cup sub-samples with errors clustered at the country-year level.

Table 3.16 presents the estimation results of Equation 3.8 for the $OIBNUM_{m,t,w}$, $OIBSH_{m,t,w}$ and $OIBDOL_{m,t,w}$ dependent variables. Similar to the stock return analysis, Table 3.16 provides some evidence in favour of the sentiment hypothesis. All statistically significant coefficients presented in Table 3.16 are of the sign predicted by the sentiment literature. Specifically, Table 3.16 documents a half-time decline in order imbalance following losses. The estimated β^{HL} coefficients for the $OIBDOL_{m,t,w}$ dependent variable are negative and statistically significant from zero. The results indicate a decrease in dollar order imbalance of 27.3% to 29.3% of a standard deviation during a half-time observation, conditional on the national team associated with market m, losing at half-time. Nonetheless, most coefficients presented in Table 3.16 are not statistically different from zero.

It is possible to exploit variation in trade characteristics to determine the trades that are most affected by sentiment. Accordingly, for each market m and World Cup sub-sample w, I dissect trades into quintiles according to dollar volume. For each quintile, I construct a unique dollar order imbalance measure. This gives five new order imbalance measures, $OIBDOL^q_{m,t,w}$ where $q \in \{1, 2, 3, 4, 5\}$. The dollar order imbalance measure for the first quintile of trades is denoted by $OIBDOL^1_{m,t,w}$ while the analogous measure for the fifth quintile is denoted by $OIBDOL^5_{m,t,w}$. The $OIBDOL^q_{m,t,w}$ variables are detrended and standardised prior to empirical investigation in the same fashion as the $OIBDOL_{m,t,w}$ variable. Thus, let $OIBDOL^q_{m,t,w}$ denote the adjusted variables.

Table 3.17 presents the Equation 3.8 estimation results with respect to the $OIBDOL^q_{m,t,w}$ dependent variables. In the interest of brevity, the results for $OIBDOL^2_{m,t,w}$ and $OIBDOL^4_{m,t,w}$ are not shown. The estimation results suggest that the half-time loss-effect identified in Table 3.16 is prevalent among trades of all sizes. For every quintile, the estimated β^{HL} coefficient ranges between -0.242 and -0.292. It is also interesting to note that Table 3.17 provides some evidence that smaller

Table 3.16
Abnormal Order Imbalance during International Football Matches

This table reports the estimation results for the following regression:

$$y_{m,t,w} = \alpha_{m,w} + \beta^{globe} r_{globe,t,w} + \beta^{HW} \mathbb{1}\{HW\}_{m,t,w} + \beta^{HL} \mathbb{1}\{HL\}_{m,t,w} + \beta^{W} \mathbb{1}\{W\}_{m,t,w} + \beta^{L} \mathbb{1}\{L\}_{m,t,w} + \epsilon_{m,t,w} \tag{3.8}$$

where $y_{m,t,w}$ is the seasonally detrended and standardised dependent variable of market m at five-minute intraday time period t and World Cup sub-sample w. The $OIBNUM_{m,t,w}$ dependent variable is the number of buyer-initiated trades less the number of seller-initiated trades on market m at time t in World Cup sub-sample w. The $OIBSH_{m,t,w}$ dependent variable is the number of buyer-initiated shares purchased less the number the seller-initiated shares sold and the $OIBDOL_{m,t,w}$ dependent variable is the buyer-initiated dollars (or local currency) paid less the seller-initiated dollars (or local currency) received, both for market m, time t and World Cup sub-sample w. Buyer- and seller-initiated trades are determined using the Lee and Ready (1991) algorithm. The $r_{globe,t,w}$ variable is the seasonally detrended five-minute return on the S&P Global Broad Market Index at time t during World Cup sub-sample w. The indicator variable, $\mathbb{1}\{HW\}_{m,t,w}$, takes the value of 1 during half-time of a football match in World Cup sub-sample w and in which the national football team associated with market m is participating, conditional on that team winning at half-time. The indicator variables, $\mathbb{1}\{HL\}_{m,t,w}$, takes the value of 1 during half-time of a football match in World Cup sub-sample w in which the national football team associated with market m is participating, conditional on that team losing at half-time. The indicator variable, $\mathbb{1}\{W\}_{m,t,w}$, takes the value of 1 between full-time of a football match in World Cup sub-sample w and the market close of market m on match day, conditional on the national football team associated with market m winning the match. The indicator variable, $\mathbb{1}\{L\}_{m,t,w}$, takes the value of 1 between full-time of a football match in World Cup sub-sample w and the market close of market m on match day, conditional on the national football team associated with market m losing the match. The t-statistics are reported in italics. The standard errors are clustered at the country-World Cup sub-sample level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively. R^2_W is a weighted coefficient of determination that only utilises observations for which at least one match outcome indicator variable is nonzero.

Dependent Variable	$OIBNUM_{m,t,w}$		$OIBSH_{m,t,w}$		$OIBDOL_{m,t,w}$	
	(1)	(2)	(1)	(2)	(1)	(2)
β^{globe}		0.081		0.060		0.076
		1.43		*1.20*		*1.38*
β^{HW}	0.063	0.053	0.003	0.047	0.033	0.042
	0.30	*0.25*	*0.02*	*0.28*	*0.22*	*0.25*
β^{HL}	-0.265*	-0.256	-0.168	-0.143	-0.273**	-0.293*
	-1.76	*-1.50*	*-1.57*	*-1.20*	*-1.97*	*-1.86*
β^{W}	0.013	0.004	0.022	0.026	-0.027	-0.029
	0.29	*0.08*	*0.51*	*0.55*	*-0.72*	*-0.72*
β^{L}	0.068	0.074	-0.040	-0.053	-0.019	-0.037
	1.01	*1.01*	*-0.62*	*-0.79*	*-0.31*	*-0.59*
Observations	273736	243981	273736	243981	273736	243981
R^2_W (%)	0.29	0.30	0.11	0.15	0.16	0.21

Table 3.17
Abnormal Order Imbalance by Quintiles during International Football Matches

This table reports the estimation results for the following regression:

$$y_{m,t,w} = \alpha_{m,w} + \beta^{globe} r_{globe,t,w} + \beta^{HW} \mathbb{1}\{HW\}_{m,t,w} + \beta^{HL} \mathbb{1}\{HL\}_{m,t,w} + \beta^{W} \mathbb{1}\{W\}_{m,t,w} + \beta^{L} \mathbb{1}\{L\}_{m,t,w} + \epsilon_{m,t,w} \tag{3.8}$$

where $y_{m,t,w}$ is the seasonally detrended and standardised dependent variable of market m at five-minute intra-day time period t and World Cup sub-sample w. The dependent variables are constructed by dissecting trades into quintiles according to dollar volume. For each quintile, I construct a unique dollar order imbalance measure. The dollar order imbalance measure for the qth quintile of trades is denoted by $OIBDOL^q_{m,t,w}$. The $OIBDOL^q_{m,t,w}$ measure is the buyer-initiated dollars (or local currency) paid less the seller-initiated dollars (or local currency) received, for market m, time t, World Cup sub-sample w and dollar volume quintile q. Buyer- and seller-initiated trades are determined using the Lee and Ready (1991) algorithm. The $r_{globe,t,w}$ variable is the seasonally detrended five-minute return on the S&P Global Broad Market Index at time t during World Cup sub-sample w. The indicator variable, $\mathbb{1}\{HW\}_{m,t,w}$, takes the value of 1 during half-time of a football match in World Cup sub-sample w and in which the national football team associated with market m is participating, conditional on that team winning at half-time. The indicator variables, $\mathbb{1}\{HL\}_{m,t,w}$, takes the value of 1 during half-time of a football match in World Cup sub-sample w in which the national football team associated with market m is participating, conditional on that team losing at half-time. The indicator variable, $\mathbb{1}\{W\}_{m,t,w}$, takes the value of 1 between full-time of a football match in World Cup sub-sample w and the market close of market m on match day, conditional on the national football team associated with market m winning the match. The indicator variable, $\mathbb{1}\{L\}_{m,t,w}$, takes the value of 1 between full-time of a football match in World Cup sub-sample w and the market close of market m on match day, conditional on the national football team associated with market m losing the match. The t-statistics are reported in italics. The standard errors are clustered at the country-World Cup sub-sample level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively. R^2_W is a weighted coefficient of determination that only utilises observations for which at least one match outcome indicator variable is nonzero.

Dependent Variable	$OIBDOL^1_{m,t,w}$		$OIBDOL^3_{m,t,w}$		$OIBDOL^5_{m,t,w}$	
	(1)	(2)	(1)	(2)	(1)	(2)
β^{globe}		0.001		0.026		-0.015
		0.02		*0.39*		*-0.24*
β^{HW}	0.316*	0.113	0.290	0.311	0.160	0.176
	1.74	*0.76*	*1.37*	*1.39*	*1.02*	*1.00*
β^{HL}	-0.258*	-0.271*	-0.292*	-0.281	-0.242*	-0.266
	-1.85	*-1.72*	*-1.87*	*-1.60*	*-1.68*	*-1.64*
β^{W}	0.012	-0.005	0.031	-0.013	-0.095	-0.110
	0.14	*-0.06*	*0.47*	*-0.21*	*-1.28*	*-1.42*
β^{L}	-0.077	-0.057	0.108	0.129	-0.030	-0.037
	-0.54	*-0.38*	*0.97*	*1.05*	*-0.34*	*-0.39*
Observations	225683	198445	225683	198445	225683	198445
R^2_W (%)	0.61	0.31	0.71	0.79	0.75	0.92

trades are more susceptible to sentiment effects. For the $OIBDOL^1_{m,t,w}$ dependent variable and first specification with a larger sample period, all estimated β^{HW}, β^{HL}, β^{W} and β^{L} coefficients are of the sign predicted by the sentiment literature. Further, the estimated β^{HW} coefficient is 0.316 and statistically different from zero at the 90% confidence level. Thus, following a winning half-time match outcome smaller trades experience a 31.6% of a standard deviation increase in dollar order imbalance.

3.4.3.2 Quote Revisions Analysis

The previous sub-section finds evidence that trade-initiators are impacted by sentiment shocks; however, it could be the case that liquidity providers are also impacted by sentiment shocks. Liquidity providers may be directly or indirectly influenced by investor sentiment. The direct channel involves sentiment-inducing events impacting on their levels of optimism. An indirect channel could involve a flow-on effect through trade initiators. That is, if trade initiators are influenced by sentiment shocks, their abnormal order flow may be mistakenly interpreted by liquidity providers as informative. Accordingly, liquidity providers may adjust their quotes in reaction to sentiment-induced order flow. If this is the case, quote revisions may be correlated to football match outcomes. Accordingly, I examine quote revisions that are not the result of price impact. The variable, $BIDREVS_{m,t,w}$, is the number of upwards bid revisions minus the number of downward bid revisions that are not caused by price impact for market m and World Cup sub-sample w at time t. The variable $ASKREVS_{m,t,w}$ is similarly defined for ask revisions. Thus, similar to the order imbalance measures, $BIDREVS_{m,t,w}$ and $ASKREVS_{m,t,w}$ should be positively related to investor sentiment. The $BIDREVS_{m,t,w}$ and $ASKREVS_{m,t,w}$ variables are standardised and detrended in the same fashion as the order imbalance measures presented in Sub-section 3.4.3.1.

To test for abnormal quoting activity surrounding World Cup football matches, I estimate Equation 3.8 with respect to $BIDREVS_{m,t,w}$ and $ASKREVS_{m,t,w}$. Table 3.18 demonstrates that the β^{HL} estimated coefficients are negative and statistically significant for the $BIDREVS_{m,t,w}$ dependent variable. This is indicative of liquidity providers revising bid-quotes downwards following losing half-time match outcomes. The bid revisions variable, $BIDREVS_{m,t,w}$, decreases by 19.1% to 24.7% of a standard deviation during losing half-time observations. Interestingly, Table 3.18 indicates that there is no significant contemporaneous decline in the ask revisions variable, $ASKREVS_{m,t,w}$. This suggests that liquidity providers on average simultaneously decrease bid-quotes and increase spreads during losing half-time periods. Table 3.16 also indicates that the half-time losing time period is the most active period for declines in order imbalance. That is, trade initiators predominantly sell shares during losing half-time match outcomes. Given that liquidity providers appear to revise bid-quotes downwards and increase spreads, the evidence suggests that liquidity providers mistakenly interpret sentiment-induced order flow as informative, resulting in a revision of the mid-quote price and an increase to the bid-ask spread, in reaction to a perceived risk of adverse selection. Thus, their is evidence that quoting behaviour can be indirectly influenced by

Table 3.18
Abnormal Quote Revisions during International Football Matches

This table reports the estimation results for the following regression:

$$\begin{aligned} y_{m,t,w} =& \alpha_{m,w} + \beta^{globe} r_{globe,t,w} + \beta^{HW} \mathbb{1}\{HW\}_{m,t,w} + \beta^{HL} \mathbb{1}\{HL\}_{m,t,w} \\ &+ \beta^{W} \mathbb{1}\{W\}_{m,t,w} + \beta^{L} \mathbb{1}\{L\}_{m,t,w} + \epsilon_{m,t,w} \end{aligned} \quad (3.8)$$

where $y_{m,t,w}$ is the seasonally detrended and standardised dependent variable of market m at five-minute intra-day time period t and World Cup sub-sample w. The $BIDREVS_{m,t,w}$ dependent variable is the number of upwards bid revisions minus the number of downward bid revisions that are not caused by price impact for market m and World Cup sub-sample w at time t. The variable $ASKREVS_{m,t,w}$ is similarly defined for ask revisions. The $r_{globe,t,w}$ variable is the seasonally detrended five-minute return on the S&P Global Broad Market Index at time t during World Cup sub-sample w. The indicator variable, $\mathbb{1}\{HW\}_{m,t,w}$, takes the value of 1 during half-time of a football match in World Cup sub-sample w and in which the national football team associated with market m is participating, conditional on that team winning at half-time. The indicator variables, $\mathbb{1}\{HL\}_{m,t,w}$, takes the value of 1 during half-time of a football match in World Cup sub-sample w in which the national football team associated with market m is participating, conditional on that team losing at half-time. The indicator variable, $\mathbb{1}\{W\}_{m,t,w}$, takes the value of 1 between full-time of a football match in World Cup sub-sample w and the market close of market m on match day, conditional on the national football team associated with market m winning the match. The indicator variable, $\mathbb{1}\{L\}_{m,t,w}$, takes the value of 1 between full-time of a football match in World Cup sub-sample w and the market close of market m on match day, conditional on the national football team associated with market m losing the match. The t-statistics are reported in italics. The standard errors are clustered at the country-World Cup sub-sample level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively. R^2_W is a weighted coefficient of determination that only utilises observations for which at least one match outcome indicator variable is nonzero.

Dependent Variable	$BIDREVS_{m,t,w}$		$ASKREVS_{m,t,w}$	
	(1)	(2)	(1)	(2)
β^{globe}		-0.064		0.153***
		-0.97		*3.10*
β^{HW}	-0.044	-0.077	0.097	0.090
	-0.23	*-0.37*	*0.87*	*0.73*
β^{HL}	-0.191*	-0.247**	-0.096	-0.124
	-1.86	*-2.35*	*-0.72*	*-0.82*
β^{W}	0.041	0.028	0.023	0.019
	0.70	*0.44*	*0.35*	*0.28*
β^{L}	0.154	0.161	0.091	0.100
	1.21	*1.20*	*1.13*	*1.22*
Observations	220780	192804	220780	192804
R^2_W (%)	0.89	1.00	0.37	0.45

investor sentiment.

3.5. Conclusion

This paper begins by re-examining the findings of the influential Edmans et al. (2007) study for an updated sample of stock market data. I consider all World Cup match outcomes up until 2014 to determine that the Edmans et al. (2007) loss-effect following negative sentiment shocks is still prevalent using the larger sample period. The sentiment effect is strongest among the sample of countries used in the Edmans et al. (2007) study and from 1973 to 1994. The second and third portions of the paper concentrate on a sub-sample of World Cup football matches that occur during the trading hours of the participating countries of each match. The analysis reveals that countries experience positive abnormal stock returns between full-time of a football match and market close, following a victory. Further, unexpected victories and victories over traditional rivals have a positive and significant marginal impact on abnormal intra-day stock returns. Thus, their is evidence of a predominant overnight loss-effect following negative sentiment shocks and an intra-day win-effect following positive sentiment shocks. This result provides support for a contemporaneous study, Cai et al. (2018), that argues that a proportion of the overnight loss-effect can be attributed to the physiological impacts of investors disrupting their regular sleeping patterns to monitor football matches.

The third portion of the paper makes use of trade and quote data to test an implicit assumption of the sentiment literature. The assumption is that sentiment influences investors' levels of optimism, which in turn influences their buying and selling behaviour. The sentiment literature often indirectly tests this assumption by observing abnormal stock returns. This paper examines whether abnormal buying and selling behaviour can be explained by football match outcomes that coincide with trading hours. Following a losing half-time match outcome, markets exhibit a 27.3% to 29.3% of a standard deviation decline in abnormal dollar order imbalance. Moreover, for small trades, there is some evidence that dollar order imbalance increases following winning half-time match outcomes. Thus, there is significant evidence in favour of the sentiment literature assumptions.

Finally, this paper examines whether liquidity providers may be influenced by sentiment. This is achieved by isolating quote revisions that are not the result of immediate price impact. The analysis reveals that downward bid-quote revisions are more likely to occur following losing half-time match outcomes. On the other hand, there is no significant impact on ask-quote revisions. The empirical results suggest that liquidity providers react to sentiment-driven selling behaviour by reducing mid-quotes and increasing the bid-ask spread. Thus, the evidence suggests that liquidity suppliers are indirectly impacted by investor sentiment and interpret sentimental investors' abnormal order flow as potentially informative.

3.6. Appendix: Placebo Tests

In this appendix, I test a number of the key findings of the study by performing placebo analyses. The placebo tests are performed by constructing counterfactual match outcomes. Section 3.6.1 tests the daily stock market returns results presented in Table 3.8. Section 3.6.2 tests the intra-day stock market return results presented in Table 3.11. Section 3.6.3 tests the order imbalance results presented in Table 3.16.

3.6.1. Placebo Test of the Daily Stock Market Return Results

The first placebo test relates to the Equation 3.2 estimation results presented in Table 3.8. The placebo test is performed by repeatedly estimating Equation 3.2 for a randomly generated set of football match outcomes. I utilise the empirical distribution of goals scored per match for each World Cup iteration to construct the counter-factual match outcomes. Thus, I use five goal distributions corresponding to each World Cup iteration. For each counter-factual match observation, I randomly draw the number of goals for the national team of interest and the opposition team from the relevant empirical goal distribution. After estimating Equation 3.2 for the counter-factual match observations, I store the estimated regression coefficients. This process is repeated 1000 times to arrive at a distribution of estimated beta coefficients. Finally, the true beta coefficients are tested for statistical significance against the counter-factual coefficient distributions.

The placebo results for the daily stock market return analysis are presented in Table 3.19. Table 3.19 demonstrates that the daily stock market return results presented in Table 3.8 are robust to the placebo test. In fact, every significant coefficient presented in Table 3.8 remains statistically significant in Table 3.19. Further, the estimated β^{EL} coefficient for the 1973 to 2014 sample period is found to be significantly different from zero by the placebo test, despite being insignificantly different from zero in Table 3.8.

3.6.2. Placebo Test of the Intra-Day Stock Market Return Results

I utilise the same procedure discussed in the previous sub-section, 3.6.1, to conduct the placebo test for the intra-day stock market return results. I perform a placebo test for the Equation 3.4 results presented in Table 3.11. The results of the intra-day stock market return placebo test are presented in Table 3.20. Importantly, every significant coefficient presented in Table 3.20 has a sign that is consistent with the sentiment predictions. Further, the counter-intuitive results presented in Table 3.11 are not present in Table 3.20. In particular, the t-statistic of -1162.77 for β^{EHL} and specification (3) in Table 3.11 is found to be -2.50 by the placebo test. Recall that the t-statistic of -1162.77 in Table 3.11 was driven by a single Elimination Stage half-time winning observation. Further, the counter-intuitive positive estimated β^{EL} coefficients presented Table 3.11 are shown to be insignificantly different from zero by the placebo test.

Table 3.19
Placebo Test of the Abnormal Daily Stock Market Return Results

This table reports the placebo test results for the Equation 3.2 estimation presented in Table 3.8. The equation of interest is:

$$\epsilon_{m,d} = \alpha + \beta^{W}\mathbb{1}\{W\}_{m,d} + \beta^{L}\mathbb{1}\{L\}_{m,d} + \beta^{GW}\mathbb{1}\{GW\}_{m,d} + \beta^{GL}\mathbb{1}\{GL\}_{m,d} + \beta^{EW}\mathbb{1}\{EW\}_{m,d} + \beta^{EL}\mathbb{1}\{EL\}_{m,d} + u_{m,d} \tag{3.2}$$

where $\mathbb{1}\{W\}_{m,d}$ is an indicator variable that takes the value of one if country m wins a FIFA World Cup football match on a day that makes d the first trading day after the match and zero otherwise and $\mathbb{1}\{L\}_{m,d}$ is an indicator variable analogously defined for losses. The indicator variable, $\mathbb{1}\{GW\}_{m,d}$, takes the value of one if country m wins a Group Stage FIFA World Cup football match on a day that makes d the first trading day after the match and zero otherwise, while $\mathbb{1}\{EW\}_{m,d}$ is analogously defined for Elimination Stage matches. The indicator variable, $\mathbb{1}\{GL\}_{m,d}$, takes the value of one if country m loses a Group Stage FIFA World Cup football match on a day that makes d the first trading day after the match and zero otherwise, while $\mathbb{1}\{EL\}_{m,d}$ is analogously defined for Elimination Stage matches. The dependent variable, $\epsilon_{m,d}$, is defined by the following equation:

$$r_{m,d} = \alpha_m + \beta^1 r_{m,d-1} + \beta^2 r_{globe,d-1} + \beta^3 r_{globe,d} + \beta^4 r_{globe,d+1} + \beta^5 D_t + \beta^6 Q_{t-1} + \epsilon_{m,d} \tag{3.1}$$

where $r_{m,d}$ is the return on market m on day d, $r_{globe,d}$ is the daily U.S. dollar return on Datastream's world market index on day d, $D_t = \{D_{1,d}, D_{2,d}, D_{3,d}, D_{4,d}\}$ are indicator variables for Monday through to Thursday and $Q_{t-1} = \{Q_{m,d-1}, Q_{m,d-2}, Q_{m,d-3}, Q_{m,d-4}, Q_{m,d-5}\}$ where $Q_{m,d-1}$ is an indicator variables that take the value of one if country m experienced a non-weekend public holiday on day $d-1$ and zero otherwise. The placebo test is performed by repeatedly estimating Equation 3.2 for a randomly generated set of football match outcomes. I utilise the empirical distribution of goals scored per match for each World Cup iteration to construct the counter-factual match outcomes. Thus, I use five goal distributions corresponding to each World Cup iteration. Then, for each counter-factual match observation, I randomly draw the number of goals for the national team of interest and the opposition team from the relevant empirical goal distribution. After estimating Equation 3.2 for the counter-factual match observations, I store the estimated regression coefficients. This process is repeated 1000 times to arrive at a distribution of estimated beta coefficients. Finally, the t-statistics for the true beta coefficients are derived from the counter-factual coefficient distributions. The t-statistics are reported in italics. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively.

Year Range	No. of Games	β^{W}	No. of Games	β^{L}	No. of Games	β^{GW}	No. of Games	β^{GL}	No. of Games	β^{EW}	No. of Games	β^{EL}
1973-2014	320	0.014	248	-0.171***								
		0.22		*-2.60*								
					200	-0.047	147	-0.164*	120	0.116	101	-0.181*
						-0.55		*-1.92*		*1.18*		*-1.83*
1973-2004	185	0.035	146	-0.330***								
		0.41		*-3.70*								
					108	0.008	84	-0.325***	77	0.073	62	-0.338**
						0.07		*-2.95*		*0.51*		*-2.22*
1973-1994	106	0.097	79	-0.345***								
		0.90		*-3.01*								
					57	0.081	43	-0.329**	49	0.117	36	-0.364**
						0.53		*-2.14*		*0.73*		*-2.29*
1995-2014	214	-0.026	169	-0.089								
		-0.30		*-1.06*								
					143	-0.096	104	-0.095	71	0.116	65	-0.079
						-0.90		*-0.86*		*0.84*		*-0.55*

Table 3.20
Placebo Test of the Abnormal Intra-Day Stock Market Return Results

This table reports the placebo test results for the Equation 3.4 estimation presented in Table 3.11. The equation of interest is:

$$r_{m,t} = \alpha_m + \beta^{globe} r_{globe,t} + \beta^{HW}\mathbb{1}\{HW\}_{m,t} + \beta^{HL}\mathbb{1}\{HL\}_{m,t} + \beta^{GHW}\mathbb{1}\{GHW\}_{m,t} + \beta^{GHL}\mathbb{1}\{GHL\}_{m,t} + \beta^{W}\mathbb{1}\{W\}_{m,t} + \beta^{L}\mathbb{1}\{L\}_{m,t} + \beta^{GW}\mathbb{1}\{GW\}_{m,t} + \beta^{GL}\mathbb{1}\{GL\}_{m,t} + \beta^{EW}\mathbb{1}\{EW\}_{m,t} + \beta^{EL}\mathbb{1}\{EL\}_{m,t} + \epsilon_{m,t} \tag{3.4}$$

where $r_{m,t}$ is the seasonally detrended five-minute intra-day return of market m at time t and $r_{globe,t}$ is the seasonally detrended five-minute return on the S&P Global Broad Market Index at time t. The indicator variable, $\mathbb{1}\{HW\}_{m,t}$, takes the value of one during half-time of a football match in which the national football team associated with market m is participating, conditional on that team winning at half-time. The indicator variables, $\mathbb{1}\{GHW\}_{m,t}$ and $\mathbb{1}\{EHW\}_{m,t}$, are the equivalents of $\mathbb{1}\{HW\}_{m,t}$ with respect to Group Stage matches and Elimination Stage matches, respectively. The indicator variables, $\mathbb{1}\{HL\}_{m,t}$, takes the value of one during half-time of a football match in which the national football team associated with market m is participating, conditional on that team losing at half-time. The indicator variables, $\mathbb{1}\{GHL\}_{m,t}$ and $\mathbb{1}\{EHL\}_{m,t}$, are the equivalents of $\mathbb{1}\{HL\}_{m,t}$ with respect to Group Stage matches and Elimination Stage matches, respectively. The indicator variable, $\mathbb{1}\{W\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m winning the match. The indicator variables, $\mathbb{1}\{GW\}_{m,t}$ and $\mathbb{1}\{EW\}_{m,t}$, are the equivalents of $\mathbb{1}\{W\}_{m,t}$ with respect to Group Stage matches and Elimination Stage matches, respectively. The indicator variable, $\mathbb{1}\{L\}_{m,t}$, takes the value of one between full-time of a football match and the market close of market m on match day, conditional on the national football team associated with market m losing the match. The indicator variables, $\mathbb{1}\{GL\}_{m,t}$ and $\mathbb{1}\{EL\}_{m,t}$, are the equivalents of $\mathbb{1}\{W\}_{m,t}$ with respect to Group Stage matches and Elimination Stage matches, respectively. The placebo test is performed by repeatedly estimating Equation 3.4 for a randomly generated set of football match outcomes. I utilise the empirical distribution of goals scored per match for each World Cup iteration to construct the counter-factual match outcomes. Thus, I use five goal distributions corresponding to each World Cup iteration. Then, for each counter-factual match observation, I randomly draw the number of goals for the national team of interest and the opposition team from the relevant empirical goal distribution. After estimating Equation 3.4 for the counter-factual match observations, I store the estimated regression coefficients. This process is repeated 1000 times to arrive at a distribution of estimated beta coefficients. Finally, the t-statistics for the true beta coefficients are derived from the counter-factual coefficient distributions. The t-statistics are reported in italics. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively.

	(1)	(2)	(3)	(4)
β^{globe}			0.056***	0.056***
			2.96	*2.96*
β^{HW}		0.032		0.037
		1.64		*1.63*
β^{HL}		-0.007		-0.008
		-0.60		*-0.62*
β^{GHW}	0.021*		0.020	
	1.86		*1.58*	
β^{GHL}	-0.001		-0.001	
	-0.10		*-0.08*	
β^{EHW}	0.071*		0.089*	
	1.70		*1.70*	
β^{EHL}	-0.064		-0.130**	
	-1.49		*-2.50*	
β^{W}	0.005		0.006*	
	1.43		*1.73*	

Table 3.20
(continued)

	(1)	(2)	(3)	(4)
β^{L}	0.002		0.001	
	0.58		*0.19*	
β^{GW}		0.002		0.003
		0.45		*0.66*
β^{GL}		-0.002		-0.003
		-0.47		*-0.72*
β^{EW}		0.015*		0.018
		1.88		*1.64*
β^{EL}		0.011		0.012
		1.35		*1.10*
Observations	7144776	7144776	6360781	6360781

3.6.3. Placebo Test of the Order Imbalance Results

I utilise the same placebo test procedure as described in Sub-section 3.6.1 to examine the validity of the abnormal order imbalance results presented in Table 3.16. The placebo test results are presented in Table 3.21. The t-statistics derived from the placebo test are remarkably similar to those of the original analysis. All significant coefficients presented in Table 3.16 remain statistically significant in Table 3.21. Thus, the abnormal order imbalance results presented in Table 3.16 are robust.

Table 3.21
Placebo Test of Abnormal Order Imbalance Results

This table reports the placebo test results for the Equation 3.8 estimation presented in Table 3.16. The equation of interest is:

$$
\begin{aligned}
y_{m,t,w} =& \alpha_{m,w} + \beta^{globe} r_{globe,t,w} + \beta^{HW} \mathbb{1}\{HW\}_{m,t,w} + \beta^{HL} \mathbb{1}\{HL\}_{m,t,w} \\
&+ \beta^{W} \mathbb{1}\{W\}_{m,t,w} + \beta^{L} \mathbb{1}\{L\}_{m,t,w} + +\epsilon_{m,t,w} \qquad (3.8)
\end{aligned}
$$

where $y_{m,t,w}$ is the seasonally detrended and standardised dependent variable of market m at five-minute intra-day time period t and World Cup sub-sample w. The $OIBNUM_{m,t,w}$ dependent variable is the number of buyer-initiated trades less the number of seller-initiated trades on market m at time t in World Cup sub-sample w. The $OIBSH_{m,t,w}$ dependent variable is the number of buyer-initiated shares purchased less the number the seller-initiated shares sold and the $OIBDOL_{m,t,w}$ dependent variable is the buyer-initiated dollars (or local currency) paid less the seller-initiated dollars (or local currency) received, both for market m, time t and World Cup sub-sample w. Buyer- and seller-initiated trades are determined using the Lee and Ready (1991) algorithm. The $r_{globe,t,w}$ variable is the seasonally detrended five-minute return on the S&P Global Broad Market Index at time t during World Cup sub-sample w. The indicator variable, $\mathbb{1}\{HW\}_{m,t,w}$, takes the value of 1 during half-time of a football match in World Cup sub-sample w and in which the national football team associated with market m is participating, conditional on that team winning at half-time. The indicator variables, $\mathbb{1}\{HL\}_{m,t,w}$, takes the value of 1 during half-time of a football match in World Cup sub-sample w in which the national football team associated with market m is participating, conditional on that team losing at half-time. The indicator variable, $\mathbb{1}\{W\}_{m,t,w}$, takes the value of 1 between full-time of a football match in World Cup sub-sample w and the market close of market m on match day, conditional on the national football team associated with market m winning the match. The indicator variable, $\mathbb{1}\{L\}_{m,t,w}$, takes the value of 1 between full-time of a football match in World Cup sub-sample w and the market close of market m on match day, conditional on the national football team associated with market m losing the match. The placebo test is performed by repeatedly estimating Equation 3.8 for a randomly generated set of football match outcomes. I utilise the empirical distribution of goals scored per match for each World Cup iteration to construct the counter-factual match outcomes. Thus, I use five goal distributions corresponding to each World Cup iteration. Then, for each counter-factual match observation, I randomly draw the number of goals for the national team of interest and the opposition team from the relevant empirical goal distribution. After estimating Equation 3.8 for the counter-factual match observations, I store the estimated regression coefficients. This process is repeated 1000 times to arrive at a distribution of estimated beta coefficients. Finally, the t-statistics for the true beta coefficients are derived from the counterfactual coefficient distributions. The t-statistics are reported in italics. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively.

Dependent Variable	$OIBNUM_{m,t,w}$		$OIBSH_{m,t,w}$		$OIBDOL_{m,t,w}$	
	(1)	(2)	(1)	(2)	(1)	(2)
β^{globe}		0.081***		0.060***		0.076***
		271.41		*173.59*		*247.68*
β^{HW}	0.063	0.053	0.003	0.047	0.033	0.042
	0.42	*0.34*	*0.02*	*0.23*	*0.23*	*0.26*

Table 3.21
(continued)

β^{HL}	-0.265*	-0.256	-0.168	-0.143	-0.273*	-0.293*
	-1.76	*-1.64*	*-0.95*	*-0.73*	*-1.91*	*-1.86*
β^{W}	0.013	0.004	0.022	0.026	-0.027	-0.029
	0.18	*0.05*	*0.27*	*0.31*	*-0.35*	*-0.37*
β^{L}	0.068	0.074	-0.040	-0.053	-0.019	-0.037
	1.00	*1.05*	*-0.49*	*-0.63*	*-0.25*	*-0.48*
Observations	273736	243981	273736	243981	273736	243981

4. Concluding Remarks

This dissertation examines liquidity and investment sentiment in financial markets. Chapter 1 "Asymmetric Liquidity Persistence" provides a comprehensive analysis of the time-series properties of liquidity. Chapter 1 demonstrates that market liquidity can be characterised as a long-memory process. Moreover, the persistence of liquidity is conditional on past market states, represented by stock market returns. Large negative returns cause liquidity persistence to initially decrease and then increase in the long-run. Thus, the persistence of liquidity is asymmetric. To account for asymmetric liquidity persistence, Chapter 1 proposes the threshold heterogeneous autoregressive (THAR) model for estimating the liquidity process. The THAR model provides for more accurate in-sample and out-of-sample estimations of liquidity. Chapter 1 concludes by testing the Amihud (2002) hypothesis of expected and unexpected liquidity. Under the assumptions of the Amihud (2002) hypotheses, the THAR model can accurately filter liquidity into its expected and unexpected components.

Chapter 2 "Discretionary Trading Surrounding Anticipated Distraction Events: the Case of the FIFA World Cup" examines discretionary liquidity trading surrounding distraction events. The analysis takes advantage of World Cup football matches that occur during trading hours to identify distraction events and shocks to the opportunity cost of monitoring markets. The analysis shows that World Cup football matches induce a decline in contemporaneous trading activity, as well as an increase in trading activity prior to match time. This is evidence of discretionary trading behaviour, whereby investors fulfil their trading demand prior to match time. The analysis concludes by demonstrating that the liquidity, volatility and price discovery dynamics on match days are consistent with the Admati and Pfleiderer (1988) discretionary trading predictions.

Chapter 3 "Sports Sentiment and Stock Returns: An Intra-day Study", utilises a similar dataset to Chapter 2 to test for investor sentiment effects surrounding football matches within 19 stock markets. Chapter 3 provides an array of empirical results that are consistent with the investor sentiment predictions. For example, winning full-time outcomes are associated with positive abnormal stock returns for the remainder of the trading day; while, unexpected victories and victories over traditional rivals have a positive and significant marginal impact on abnormal stock returns. Chapter 3 extends the investor sentiment literature by analysing trade and quote data to gain an insight into the underlying mechanism of sentiment effects. Chapter 3 demonstrates that football match outcomes can predict abnormal buying and selling behaviour. Further, liquidity suppliers appear to interpret this abnormal buying and selling behaviour as potentially informative. The evidence suggests that, in reaction to abnormal order imbalance, liquidity providers adjust mid-quotes and increase bid-ask spreads to counteract potential adverse selection costs.

The final chapter of this thesis provides some further discussion regarding the analyses presented in chapters 1, 2 and3 and outlines some opportunities for further research.

5. Limitations and Future Research

This chapter outlines opportunities for further research with respect to the three major dissertation chapters:

- Chapter 1 Asymmetric Liquidity Persistence;
- Chapter 2 Discretionary Trading Surrounding Anticipated Distraction Events: the Case of the FIFA World Cup; and,
- Chapter 3 Sports Sentiment and Stock Returns: An Intra-day Study.

5.1. Asymmetric Liquidity Persistence - Limitations and Future Research

The following sub-sections discuss limitations and opportunities for further research with respect to Chapter 1 "Asymmetric Liquidity Persistence". Sub-section 5.2.1 discusses recent developments in the illiquidity-asset-pricing literature that have limiting implications for Chapter 1. Sub-section 5.2.2 outlines some theories of long-memory that could be incorporated into Chapter 1. Sub-section 5.2.3 discusses the opportunity to include alternative illiquidity measures in the analysis.

5.1.1. Recent Developments in the Illiquidity-Asset-Pricing Literature

A major limitation for further research with respect to Chapter 1 relates to recent developments in the illiquidity-asset-pricing literature. Developments in this literature have been driven by a forthcoming issue of the *Critical Finance Review* titled *Liquidity: Replications, Extensions, and Critique*. This issue seeks to replicate and verify findings in some of the most influential papers within this literature: Amihud (2002); Pástor and Stambaugh (2003); and, Acharya and Pedersen (2005). With respect to the Amihud (2002) study, the results of the Drienko et al. (2018) and Harris and Amato (2018) replication studies are similar to those presented in Section 1.6 *Amihud (2002) Hypotheses*. Section 1.6 finds evidence in favour of Amihud's (2002) second hypothesis and limited evidence in favour of his first. Drienko et al. (2018) attribute this to "a decline in the sensitivity of investors to illiquidity risk over the last two decades, a period during which technological innovations and decimalization have markedly reduced transaction costs and increased stock liquidity".

In two new replication and extension studies, Eiichiro Kazumori and Yu (2018) and Holden and Nam (2018) evaluate the validity of the Acharya and Pedersen (2005) Liquidity-adjusted Capital Asset Pricing Model (LCAPM). Holden and Nam (2018) find that the predictions of the LCAPM do not hold for a more recent sample period of 2000 to 2015. Eiichiro Kazumori and Yu (2018) test the "net beta" LCAPM of Acharya and Pedersen (2005) and find that the net-beta LCAPM has a failure rate of 64% in their replication tests. Eiichiro Kazumori and Yu (2018) attribute this result to severe multi-collinearity problems. For example, during the 2000 to 2006 sample period, Eiichiro Kazumori and Yu (2018) find a correlation of -0.983 between Acharya and Pederson's

(2005) second and fourth liquidity betas at the portfolio level.

The original intention of Chapter 1 was to apply the THAR(3) model to the theoretical predictions of the Amihud (2002) and Acharya and Pedersen (2005) studies. While, the THAR(3) model is applied to the theoretical predictions of Amihud (2002) in Section 1.6, the model could also be applied to the LCAPM by utilising illiquidity innovations derived from the model to calculate the illiquidity risk factors of Acharya and Pedersen (2005).[31] The motivation for this application of the THAR(3) model would be to obtain a more accurate estimation of the liquidity-adjusted market risk premium and to test whether the LCAPM is robust to a different specification of the liquidity process. This analysis intended to replicate the Acharya and Pedersen (2005) methodology for estimating the LCAPM and to subsequently compare the results to the LCAPM results derived from the THAR(3) model. Unfortunately, this avenue of research was not feasible. This is because, in results not shown, we were unable to replicate the baseline results of Acharya and Pedersen (2005). Given the recent findings of Eiichiro Kazumori and Yu (2018) and Holden and Nam (2018), it is apparent that other researchers have encountered similar difficulties.

5.1.2. Theories of Long-Memory

A potential avenue for further research could be to draw closer ties between the THAR(3) model the theoretical explanations for long-memory in market variables. For example, Corsi (2009) motivates his HAR model by citing the Heterogeneous Market Hypothesis of Müller, Dacorogna, Davé, Pictet, Olsen, and Ward (1993). Corsi (2009) postulates that long-run dependencies in volatility could be the result of heterogeneous agents and in particular, agents with different investment horizons.

In the quantitative finance literature, long-memory properties of intra-day market variables are attributed to various explanations. Bouchaud, Farmer, and Lillo (2009) divide these explanations into two classes. The first class of explanations identify long-memory as a function of individuals' order flows. The second class of explanations promote long-memory as the result of the aggregation of individuals' order flow.

The predominant theory of the first class of explanations is order-splitting or "delayed market clearing". This hypothesis is first stipulated by Bouchaud, Gefen, Potters, and Wyart (2004) and formalised in Lillo, Mike, and Farmer (2005). Evidence of order-splitting can be found in Chan and Lakonishok (1993, 1995), where approximately one half of large institutional ''package'' trades take longer than a week to clear. Kyle, Obizhaeva, and Wang (2014) assert that "asset managers often...[acquire] positions over days, weeks, or even months". Perhaps the most profound evidence of the order-splitting explanation of long-memory in liquidity is found in Gerig (2007). Using brokerage ID data for London Stock Exchange-listed stocks, Gerig (2007) shows that order flow from the same brokerage demonstrates the strongest long-memory (autocorrelation function with

[31] Acharya and Pedersen (2005) utilise AR(1) innovations of monthly illiquidity to arrive at their liquidity risk factors.

power law exponent -0.4); all order flow demonstrates very strong long-memory (autocorrelation function with power law exponent -0.67); while, order flow from different brokerages do not display any significant autocorrelation. Hence, it appears as though, order-splitting by individuals is the primary cause for long-memory in order flow.

Despite the strong evidence of the order-splitting explanation of long-memory in liquidity, there are a number of other explanations from the aggregated order flow perspective. Wyart and Bouchaud's (2007) self-referential behaviour theory claims that agents build strategies using correlations estimated from historical financial data. This leads to a feedback or self-fulfilling prophecy effect, whereby the market switches between two long-lived states. LeBaron and Yamamoto (2007) show that long-memory in trading volume, volatility and market order signs, in an efficient market, could be the result of the learning and adaptation of heterogeneous investors. Yamamoto (2011) attributes long-memory in order flow to agents placing more (less) aggressive orders when the same side of the order book is deep (thin). In addition, long-memory in volatility has been attributed to herding behaviour (Alfarano, Lux, and Wagner (2008)) and chartists's trend-following behaviour (Chiarella, Iori, and Perelló (2009)).

To test the theories of long-memory in financial markets, the analysis of Chapter 1 could be extended to examine long-memory and the performance of the THAR(3) model in different settings. For example, as investors in more illiquid stocks can be expected to have longer investment horizons (Stoll and Whaley (1983)), it could be that these stocks have greater persistence in liquidity. Further, it could be that illiquid stocks are more vulnerable to liquidity dry-ups. Thus, the benefits of incorporating a threshold component into the HAR framework might be more significant for illiquid stocks. Moreover, it could be that package trades and delayed market clearing are negatively correlated to block trades negotiated in an upstairs market. If this is the case, under the order-splitting theory of Bouchaud et al. (2004) and Lillo et al. (2005), long-memory in liquidity might be more severe for stocks without an active upstairs market. Finally, it might be possible to obtain broker ID data in a similar manner to Gerig (2007) to identify package trades in order to assess the performance of the THAR(3) model as a function of order-splitting. If such data is not available, it might be possible to identify package trades around index reconstitutions submitted by those managers of index-tracking funds looking to reduce their price impact.

5.1.3. Additional Illiquidity Measures

Another potential avenue for further research could be to include alternative illiquidy measures into the analysis of Chapter 1. As it stands, Chapter 1 only presents results concerning the modified Amihud (2002) measure of illiquidity, $ILLIQ_t$. In results, not shown, similar findings are made for the standard Amihud (2002) measure, $AMIHUD_t$. This is not surprising as the measures are closely related and highly correlated, as shown in Table 1.2. Alternative illiquidity measures could include a daily average bid-ask spread measure, dollar-weighted effective spread measure or a dollar-weighted realised spread measure. It is likely that the THAR(3) model would perform well for these processes, given the long-memory properties of the bid-ask spread, presented by Plerou, Gopikrishnan, and Stanley (2005).

5.2. Discretionary Trading Surrounding Anticipated Distraction Events: the Case of the FIFA World Cup - Limitations and Future Research

There are a number of opportunities to be explored with respect to Chapter 2. Sub-section 5.2.1 discusses opportunities to exploit country-level and stock-level cross-sectional variation. Sub-section 5.2.2 considers the USA sub-sample in isolation. Sub-section 5.2.3 discusses controlling for news events within trading hours. In Sub-section 5.2.4, I conduct a preliminary analysis of alternative high-frequency measures of liquidity.

5.2.1. Country-level and Stock-level Cross-sectional Variation

Chapter 2 utilises intra-day stock market data from 22 countries from 1998 to 2014. One opportunity for future research is to exploit time-series and cross-sectional variation afforded by the comprehensive stock market sample. For example, Table 2.5 gives an insight into possible cross-sectional variation that is yet to be explored. Table 2.5 indicates that some countries experience significant declines in trading activity for the entire trading day when football matches coincide with football matches. For example, the Argentina, Brazil, France, Germany, Greece, Ireland and Italy sub-samples all experience a statistically significant decline in daily $VOL_{m,t,w}$, $DVOL_{m,t,w}$, $TRADES_{m,t,w}$, $BDVOL_{m,t,w}$ and $BDVOL_{m,t,w}$. These countries are generally considered to have a strong football following. Thus, it may be worthwhile isolating these countries with a strong football following for further analysis.[32] Further, considering that these countries experience a significant decrease in *daily* trading activity for the entire match-day, it could be that these countries experience discretionary liquidity trading at the daily frequency, as in Foster and Viswanathan (1990). If this is the case, a positive abnormal amount of trading activity may occur on the trading days before match days.[33]

Another opportunity for further research is to concentrate on a sub-sample of liquid countries, with higher turnover, that are more likely to have elevated levels of discretionary liquidity trading. For example, Figure 5.1 compares the 1998 to 2006 sample period to the 2010 to 2014 sample period. Figure 5.1 demonstrates that the more recent sample period exhibits a greater level of abnormal trading activity in the pre-match period of the trading day. This is unsurprising as the more recent sample period can be expected to be the more liquid portion of the sample.

[32] For example, in their contemporaneous study, Cai et al. (2018) limit their analysis to 16 countries that have a successful football history. Cai et al. (2018) select countries that have finished first, second, third or fourth in a World Cup between 1982 and 2014, conditional on data availability.

[33] Edmans et al. (2007) do not find any evidence of abnormal trading volume on trading days after matches; however, their sample of football matches includes matches that occurred outside of trading hours.

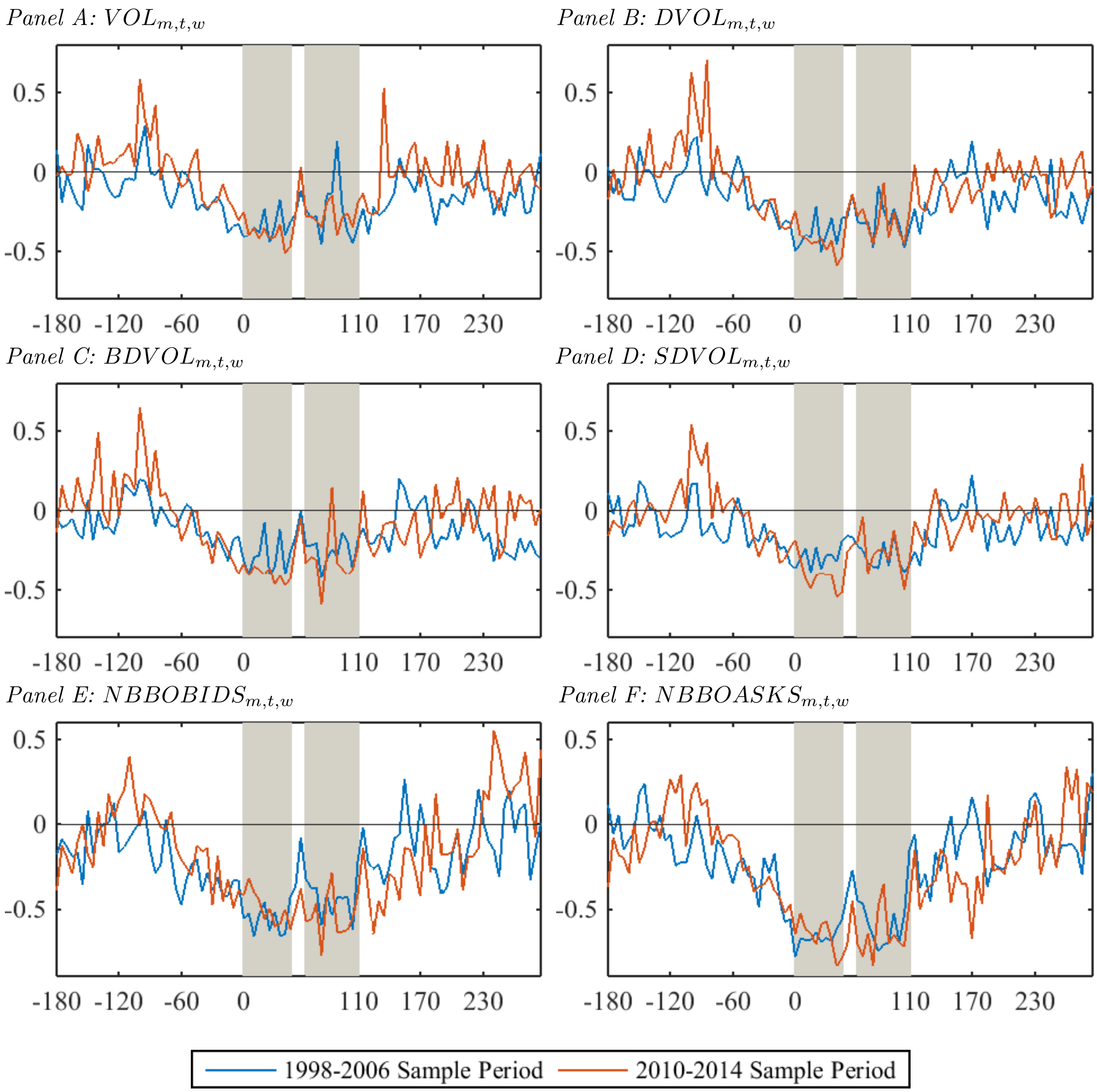


Fig. 5.1. Trading Activity on Match Days by World Cup. This figure plots the mean standardised, seasonally detrended trading activity variables on match days for the 1998 to 2006 (blue line) sample period and the 2010 to 2014 sample period (orange line). The market variables are detrended using the methodology of Gallant et al. (1992) for month-of-the-year, day-of-the-week and five-minute time-of-the-day effects. Volume for market m at time t is denoted by $VOL_{m,t,w}$, dollar volume by $DVOL_{m,t,w}$, number of trades by $TRADES_{m,t,w}$, buyer-initiated dollar trading volume by $BDVOL_{m,t,w}$, seller-initiated dollar trading volume by $SDVOL_{m,t,w}$, number of bids at the national best bid price by $NBBOBIDS_{m,t,w}$ and number of asks at the national best ask price is denoted by $NBBOASKS_{m,t,w}$. The average first-half and second-half time periods are shaded in grey. The x-axis of each plot indicates the number of minutes from kick-off time.

5.2.2. USA Sub-sample

The inclusion of the United States in the Chapter 2 stock market sample is somewhat contentious. This is because the United States is not generally considered a country with a large football following. For example, the premier American football (or soccer) league, Major League Soccer (MLS), is not defined as one of the United States' major professional sports leagues. The major sports leagues, known as the "Big Four", are defined as Major League Baseball (MLB), the National Basketball Association (NBA), the National Football League (NFL) and the National Hockey League (NHL). Given the limited popularity of football in the United States, it could be that World Cup football matches do not have a significant impact on trading activity. If this is the case, it might be more appropriate to exclude the United States from the analysis presented in Chapter 2. Indeed, Edmans et al. (2007) exclude the United States from their study for this exact reason. In light of this, this sub-section presents the trading activity and market condition results for the USA sub-sample. It should be noted; however, that the United States was only involved in seven World Cup football matches that occurred during trading hours between 1998 and 2014. Thus, I only consider this sub-sample analysis to be anecdotal, rather than scientifically rigorous.

Table 5.1 Panel A shows the Equation 2.2 estimation results for $m = USA$ and the $VOL_{USA,t,w}$, $DVOL_{USA,t,w}$, $TRADES_{USA,t,w}$, $BDVOL_{USA,t,w}$ and $SDVOL_{USA,t,w}$ dependent variables. The estimation results reveal that there is a reduction in trading activity during and in the immediate vicinity of match time in the United States. All estimated β_1, β_7, β_8, β_9 and β_{11} coefficients are negative. Further, all estimated β_τ coefficients for $\tau \in \{12, 13, 14, 15, 16\}$ in Table 5.1 Panel A are negative. Thus, trading activity is generally reduced following a World Cup football match in the United States. With respect to the pre-match period, the empirical results are mixed. That is, some estimated β_τ coefficients for $\tau \in \{2, 3, 4, 5, 6\}$ are negative and some are positive. Thus, for the USA sub-sample, their is no clear evidence of discretionary trading on match days. This is consistent with the notion that the USA does not have an overly strong football following.

Table 5.1 Panel B shows the Equation 2.2 estimation results for $m = USA$ and the $AMIHUD_{USA,t,w}$, $MAMIHUD_{USA,t,w}$, $\sigma_{USA,t,w}$ and $PD_{USA,t,w}$ dependent variables. As trading activity is not significantly increased prior to match time for the USA sub-sample, I do not form expectations regarding market conditions during this time period. Thus, Hypothesis H1 is not valid for the USA sub-sample. Nonetheless, with respect to a decrease in trading activity during match time, one might expect a contemporaneous increase in price impact costs, as well as a corresponding decrease to volatility and price discovery, as predicted by Hypothesis H2. Table 5.1 Panel B demonstrates that volatility and price discovery are generally decreased during match time but price impact costs, represented by the $MAMIHUD_{USA,t,w}$ measure, are similarly reduced during match time. Thus, the market conditions observed for the USA sub-sample are not consistent with the notion of discretionary trading behaviour, which is in turn consistent with the Table 5.1 Panel A results that do not find significant evidence of discretionary trading behaviour in the USA sub-sample. For this reason, it may be appropriate to exclude the USA sub-sample from further research, as in

Table 5.1
USA Trading Activity and Market Conditions On Match Days

This table reports the estimation results for the following regression:

$$\begin{aligned} DEP_{m,t,w} =& \alpha_0 + \beta_0 GD_{m,t,w} + \sum_{\tau=1}^{6} PRE^{K-\tau}_{m,t,w}\beta_\tau + \beta_7 D^F_{m,t,w} + \beta_8 D^H_{m,t,w} + \beta_9 D^S_{m,t,w} \\ &+ \beta_{10} D^E_{m,t,w} + \sum_{\tau=1}^{6} POST^{F+\tau}_{m,t,w}\beta_{10+\tau} + \epsilon_{m,t,w} \end{aligned} \tag{2.2}$$

where $DEP_{m,t,w}$ is the dependent variable for $m = USA$ at time t and World Cup sub-sample w. Each dependent variable is formed by pooling the standardised, detrended market-level observations. The $GD_{m,t,w}$ indicator variable is a match day indicator variable that takes the value of one if five-minute intra-day time period t coincides with a trading day in which country m is open for trading simultaneous to country m participating in a football match. The $D^F_{m,t,w}$ indicator variable takes the value of one if country m is playing in the first-half of a football match at time t in World Cup sub-sample w. The $D^S_{m,t,w}$ indicator variable is the analogous variable for the second-half of a football match. The $D^H_{m,t,w}$ indicator variable takes the value of one if it is half-time at time t for a match involving country m in World Cup sub-sample w. The $D^E_{m,t,w}$ indicator variable takes the value of one if it is extra-time at time t for a match involving country m in World Cup sub-sample w. The remaining indicator variables capture abnormal trading activity outside of match time on match days. The $PRE^{K-\tau}_{m,t,w}$ indicator variables take the value of one if t is 30τ minutes before a first-half kick-off observation or between 30τ and $30(\tau-1)$ minutes before a first-half kick-off observation, given that country m is participating in the match in World Cup sub-sample w. The $POST^{F+\tau}_{m,t,w}$ indicator variables take the value of one if t is 30τ minutes after a full-time observation or between 30τ and $30(\tau+1)$ minutes after a full-time observation, given that country m is participating in the match in World Cup sub-sample w. Volume for market m at time t and World Cup sub-sample w is denoted by $VOL_{m,t,w}$, dollar volume by $DVOL_{m,t,w}$, number of trades by $TRADES_{m,t,w}$, buyer-initiated dollar trading volume by $BDVOL_{m,t,w}$, seller-initiated dollar trading volume by $SDVOL_{m,t,w}$, the mean Amihud (2002) measure by $AMIHUD_{m,t,w}$, the mean modified Amihud (2002) measure by $MAMIHUD_{m,t,w}$, the mean volatility by $\sigma_{m,t,w}$ and the mean gross price discovery by $PD_{m,t,w}$. The t-statistics are reported in italics. The standard errors are clustered at the country level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively. R^2_W is a weighted coefficient of determination that only utilises observations for which at least one non-constant independent variable is nonzero.

Panel A: Trading Activity

$DEP_{USA,t,w}$	$VOL_{USA,t,w}$	$DVOL_{USA,t,w}$	$TRADES_{USA,t,w}$	$BDVOL_{USA,t,w}$	$SDVOL_{USA,t,w}$
$GD_{m,t,w}$	0.174	0.233	0.321	0.105	0.193
	0.64	*0.84*	*0.76*	*0.49*	*0.91*
$PRE^{K-6}_{m,t,w}$	0.089	-0.001	0.153	-0.467**	0.695***
	0.34	*-0.01*	*0.37*	*-2.18*	*3.35*
$PRE^{K-5}_{m,t,w}$	-0.199	-0.288*	-0.219	-0.267***	-0.084
	-1.31	*-1.87*	*-1.11*	*-3.28*	*-0.46*
$PRE^{K-4}_{m,t,w}$	-0.181	-0.192	-0.135	-0.140	-0.069
	-1.25	*-1.34*	*-0.56*	*-1.02*	*-0.26*
$PRE^{K-3}_{m,t,w}$	-0.079	-0.156	-0.186	-0.098	-0.141
	-0.42	*-0.79*	*-0.52*	*-0.70*	*-0.58*

Table 5.1

(continued)

$DEP_{USA,t,w}$	$VOL_{USA,t,w}$	$DVOL_{USA,t,w}$	$TRADES_{USA,t,w}$	$BDVOL_{USA,t,w}$	$SDVOL_{USA,t,w}$
$PRE^{K-2}_{m,t,w}$	-0.184	-0.226	-0.116	-0.273	-0.003
	-0.89	*-1.09*	*-0.30*	*-1.04*	*-0.01*
$PRE^{K-1}_{m,t,w}$	-0.294*	-0.363**	-0.456*	-0.216*	-0.368**
	-1.74	*-2.26*	*-1.65*	*-1.89*	*-2.15*
$D^{F}_{m,t,w}$	-0.200	-0.310	-0.512**	-0.229	-0.174**
	-0.77	*-1.29*	*-2.12*	*-0.72*	*-2.23*
$D^{H}_{m,t,w}$	-0.237***	-0.302***	-0.631***	-0.244***	-0.262**
	-5.80	*-5.26*	*-6.05*	*-3.48*	*-2.20*
$D^{S}_{m,t,w}$	-0.423*	-0.454*	-0.670	-0.308	-0.397**
	-1.66	*-1.75*	*-1.59*	*-1.41*	*-1.98*
$POST^{F+1}_{m,t,w}$	-0.575**	-0.557**	-0.882**	-0.404**	-0.474**
	-2.23	*-2.13*	*-2.22*	*-1.99*	*-2.24*
$POST^{F+2}_{m,t,w}$	-0.370	-0.437	-0.566	-0.318	-0.319
	-1.28	*-1.50*	*-1.24*	*-1.52*	*-1.15*
$POST^{F+3}_{m,t,w}$	-0.415	-0.454	-0.624	-0.329	-0.379*
	-1.53	*-1.63*	*-1.44*	*-1.56*	*-1.70*
$POST^{F+4}_{m,t,w}$	-0.233	-0.265	-0.370	-0.121	-0.172
	-0.83	*-0.97*	*-0.77*	*-0.56*	*-0.76*
$POST^{F+5}_{m,t,w}$	-0.365	-0.370	-0.375	-0.258	-0.206
	-1.38	*-1.35*	*-0.89*	*-1.16*	*-0.99*
$POST^{F+6}_{m,t,w}$	-0.300	-0.294	-0.241	-0.209	-0.282
	-0.79	*-0.75*	*-0.37*	*-0.65*	*-0.95*
Observations	19645	19647	19656	19648	19649
R^2_W (%)	5.77	6.69	8.10	3.08	9.17

Panel B: Market Conditions

$DEP_{m,t,w}$	$AMIHUD_{USA,t,w}$	$MAMIHUD_{USA,t,w}$	$\sigma_{USA,t,w}$	$PD_{USA,t,w}$
$GD_{m,t,w}$	0.040	0.468***	0.454	0.350
	1.41	*3.21*	*1.25*	*1.61*
$PRE^{K-6}_{m,t,w}$	-0.068**	-0.273**	-0.155	
	-2.36	*-2.05*	*-0.43*	
$PRE^{K-5}_{m,t,w}$	-0.058*	-0.353**	-0.102	
	-1.90	*-2.41*	*-0.46*	
$PRE^{K-4}_{m,t,w}$	-0.083	-0.155	-0.027	0.063
	-0.38	*-0.60*	*-0.12*	*0.29*
$PRE^{K-3}_{m,t,w}$	0.182	-0.385***	-0.035	-0.173
	1.29	*-2.71*	*-0.13*	*-0.68*

Table 5.1
(continued)

$DEP_{m,t,w}$	$AMIHUD_{USA,t,w}$	$MAMIHUD_{USA,t,w}$	$\sigma_{USA,t,w}$	$PD_{USA,t,w}$
$PRE^{K-2}_{m,t,w}$	-0.147	-0.459***	0.021	0.148
	-1.14	*-3.09*	*0.06*	*0.79*
$PRE^{K-1}_{m,t,w}$	0.212	-0.384***	-0.448*	-0.490**
	1.50	*-3.76*	*-1.79*	*-2.12*
$D^{F}_{m,t,w}$	-0.055	-0.183*	-0.433**	-0.344**
	-1.36	*-1.79*	*-2.42*	*-2.01*
$D^{H}_{m,t,w}$	-0.158*	-0.488	-0.633***	-1.007***
	-1.96	*-1.51*	*-5.53*	*-3.00*
$D^{S}_{m,t,w}$	0.072	-0.062	-0.573	-0.515**
	0.39	*-0.30*	*-1.62*	*-2.14*
$POST^{F+1}_{m,t,w}$	-0.087	0.123	-0.829**	-0.553**
	-0.87	*0.62*	*-2.40*	*-2.29*
$POST^{F+2,w}_{m,t}$	-0.314	-0.034	-0.542	-0.615**
	-1.48	*-0.17*	*-1.42*	*-2.07*
$POST^{F+3}_{m,t,w}$	0.022	-0.499***	-0.548	-0.781*
	0.41	*-2.59*	*-1.49*	*-1.83*
$POST^{F+4}_{m,t,w}$	-0.168	-0.200*	-0.528	-0.638*
	-1.46	*-1.76*	*-1.28*	*-1.74*
$POST^{F+5}_{m,t,w}$	-0.054	0.129	-0.169	-0.773**
	-0.94	*0.73*	*-0.45*	*-2.03*
$POST^{F+6}_{m,t,w}$	0.289***	-0.010	-0.099	-0.721***
	2.80	*-0.10*	*-0.18*	*-3.54*
Observations	18522	18785	19651	16403
R^2_W (%)	3.21	6.38	9.27	12.47

Edmans et al. (2007).

5.2.3. News Events

The analysis presented in Chapter 2 assumes that the abnormal intra-day trading pattern observed on match days is driven by the increased opportunity cost associated with monitoring markets when there is a concurrent football match. Alternatively, it could be that the abnormal trading activity observed on match days is due to abnormal information flows. While the methodology of Gallant et al. (1992) can control for normal intra-day and seasonal trends in information flow, the adjustment process of Gallant et al. (1992) cannot control for an extraordinary sequence of information events.[34] While it is not immediately obvious why exogenously determined new events would be impacted by football matches, it is plausible that information dissemination may be impacted by distraction events. Distraction events may impact investors' or news distributors' ability to process information content in a timely and efficient manner. This disruption to the information dissemination process may subdue contemporaneous trading activity and price discovery. If this is the case, one might expect an abnormal positive amount of trading activity and price discovery at the cessation of the distraction event. As this is not observed in the stock market data, it is unlikely that the abnormal market conditions observed on match days are the result of a disruption to the information dissemination process. Thus, the observed market conditions are more likely due to the Admati and Pfleiderer (1988) notion of informed traders with private information timing their trades to minimise price impact.

On the other hand, it is plausible that endogenously determined news events could be a contributing factor to the main results presented in Chapter 2. An endogenously determined news event could include any news event that is released under the timing discretion of a firm. For example, some previous studies have suggested that firms may release poor earnings announcements when traders are distracted. Patell and Wolfson (1982) demonstrate that overnight earnings announcements are more likely to followed by a negative stock market reaction; while, intra-day earnings announcements are more likely to be followed by a positive stock market reaction. This result is validated by Damodaran (1989). Moreover, the empirical evidence of Penman (1987) indicates that earnings announcements after the close-of-trade on a Friday are more likely to be considered poor earnings announcements. These studies suggest that managers release poor earnings announcements outside of trading hours in an attempt to avoid a severe devaluation of the firm's stock price. In addition, Dellavigna and Pollet (2009) demonstrate that earnings announcements released on Fridays are associated with greater post-earnings announcement drift, which the authors attribute to distraction and the limited attention of investors. If this is the case, it could be that managers might time earnings announcements to coincide with alternative distraction events. This could include football matches. If this is the case, than the flow of information on match days may be abnormal. This could comprise the analysis conducted in Chapter 2 if market sensitive information is frequently released one to two hours before football matches. Accordingly, an important avenue for further research is to control for information flows on match days. This could be achieved by incorporating RavenPack News Analytics data into the analysis.

5.2.4. Alternative Illiquidity Measures

The investigation presented in Chapter 2 utilises two measure of illiquidity; the Amihud (2002) measure, $AMIHUD_{m,t,w}$, and a modified Amihud (2002) measure, $MAMIHUD_{m,t,w}$. The Amihud (2002) measure and thereby its modified version are typically used as proxies for alternative high-frequency measures of illiquidity that are often difficult to calculate. This is because high-

[34]For example, Penman (1987) finds seasonal trends in the distribution of earnings announcements.

frequency liquidity measures are derived from processing large intra-day trade and quote datasets. Given that Chapter 2 makes use of trade and quote data, future research should include high-frequency measures of illiquidity in the main analysis. Accordingly, in this sub-section, I include some preliminary results using high-frequency illiquidity measures.

Several high-frequency illiquidity measures could be included in future research. First, a high-frequency measure of price impact could be included in the analysis. The five-minute price impact measure, $PI_{i,k}$, for stock i and trade k is defined by Goyenko, Holden, and Trzcinka (2009) as:

$$PI_{i,k} = \begin{cases} 2 \cdot (\ln(M_{i,k+5}) - \ln(M_{i,k})) & \text{if the } k\text{th trade is buy} \\ 2 \cdot (\ln(M_{i,k}) - \ln(M_{i,k+5})) & \text{if the } k\text{th trade is sell} \end{cases} \tag{5.1}$$

where $M_{i,k}$ is the mid-quote for stock i at the time at which the kth trade is executed and $M_{i,k+5}$ is the mid-quote five-minutes after the kth trade is executed for stock i. Further, Goyenko et al. (2009) define the effective spread measure, $ES_{i,k}$, for stock i and trade k as:

$$ES_{i,k} = 2 \cdot |\ln(P_{i,k}) - \ln(M_{i,k})| \tag{5.2}$$

where $P_{i,k}$ is the transaction price for trade k and stock i. Finally, Goyenko et al. (2009) define the realised spread measure, $RS_{i,k}$, for stock i and trade k as:

$$RS_{i,k} = \begin{cases} 2 \cdot (\ln(P_{i,k}) - \ln(P_{i,k+5})) & \text{if the } k\text{th trade is buy} \\ 2 \cdot (\ln(P_{i,k+5}) - \ln(P_{i,k})) & \text{if the } k\text{th trade is sell} \end{cases} \tag{5.3}$$

where $P_{i,k+5}$ is the transaction price for a trade five-minutes after trade k. Market-level high-frequency illiquidity measures can be obtained by taking a dollar-weighted average of each illiquidity value associated with each trade within each five-minute period. The market level price impact, $PI_{m,t,w}$, effective spread, $ES_{m,t,w}$, and realised spread, $RS_{m,t,w}$, variables are defined in this manner. Similar to the main analysis, the measures are standardised and detrended using the methodology of Gallant et al. (1992).

Figure 5.2 displays the mean $PI_{m,t,w}$, $ES_{m,t,w}$, and $RS_{m,t,w}$ measures on match days. The key insight from Figure 5.2 is that the high-frequency illiquidity measures do not appear to display any distinct intra-day patter over match days. Moreover, Figure 5.3 displays the mean $PI_{m,t,w}$, $ES_{m,t,w}$, and $RS_{m,t,w}$ measures on match days for developed and developing countries. Similar to the entire sample, neither the developed nor the developing sub-sample exhibits an intra-day trend in high-frequency illiquidity across match days. Table 5.2 reports the estimation results of Equation 2.2 for the $PI_{m,t,w}$, $ES_{m,t,w}$, and $RS_{m,t,w}$ dependent variables. Consistent with figures 5.2 and 5.3, the estimation results presented in Table 5.2 do not reveal a significant high-frequency liquidity intra-day trend on match days.

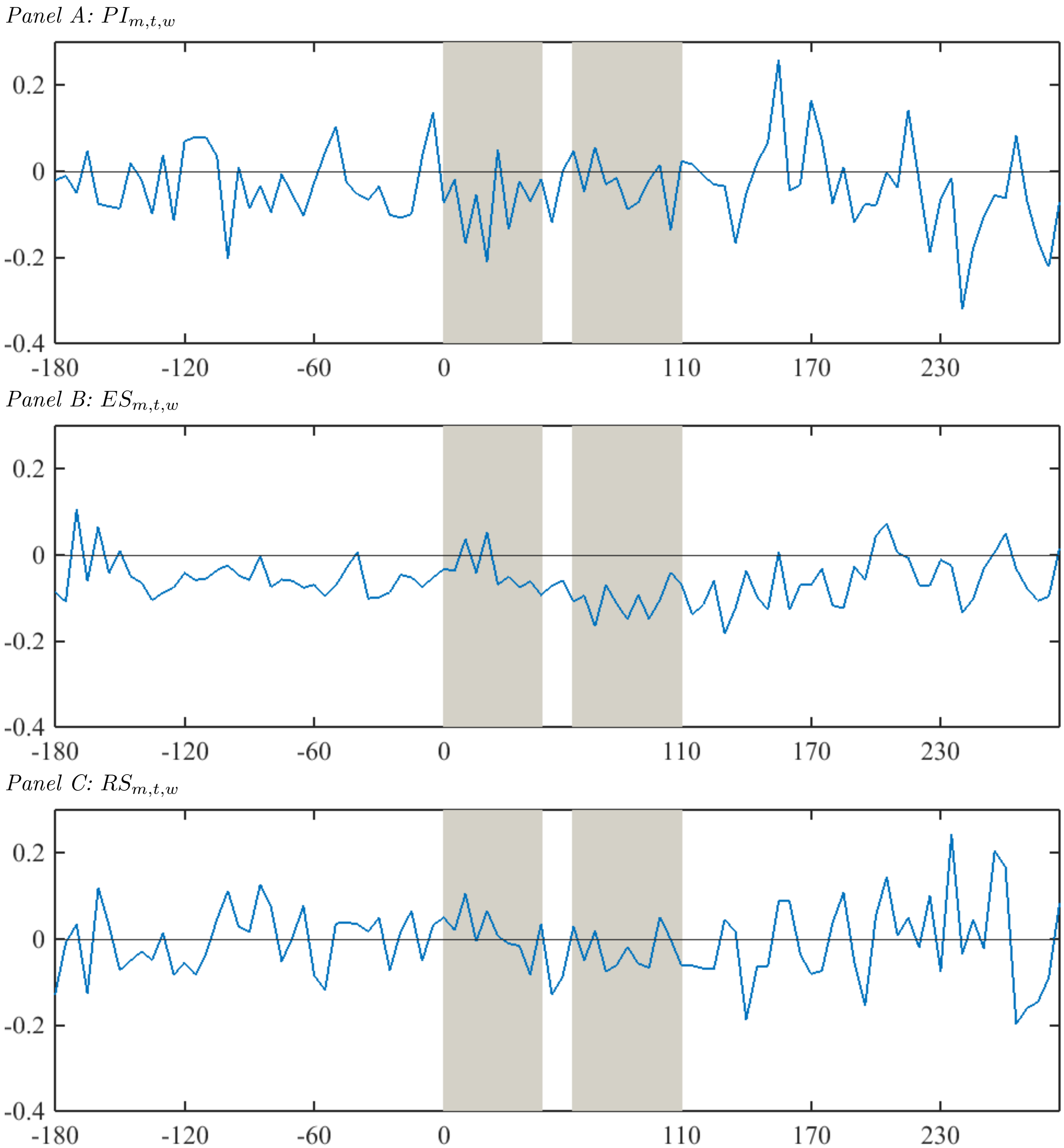


Fig. 5.2. High-Frequency Illiquidity Measures During Match Days. This figure plots the mean standardised, seasonally detrended high-frequency liquidity variables on match days. The market variables are detrended using the methodology of Gallant et al. (1992) for month-of-the-year, day-of-the-week and time-of-the-day effects. Five-minute price impact for market m at time t and World Cup sub-sample w is denoted by $PI_{m,t,w}$, effective spread by $ES_{m,t,w}$ and realised spread by $RS_{m,t,w}$. The average first-half and second-half time periods are shaded in grey. The x-axis of each plot indicates the number of minutes from kick-off time.

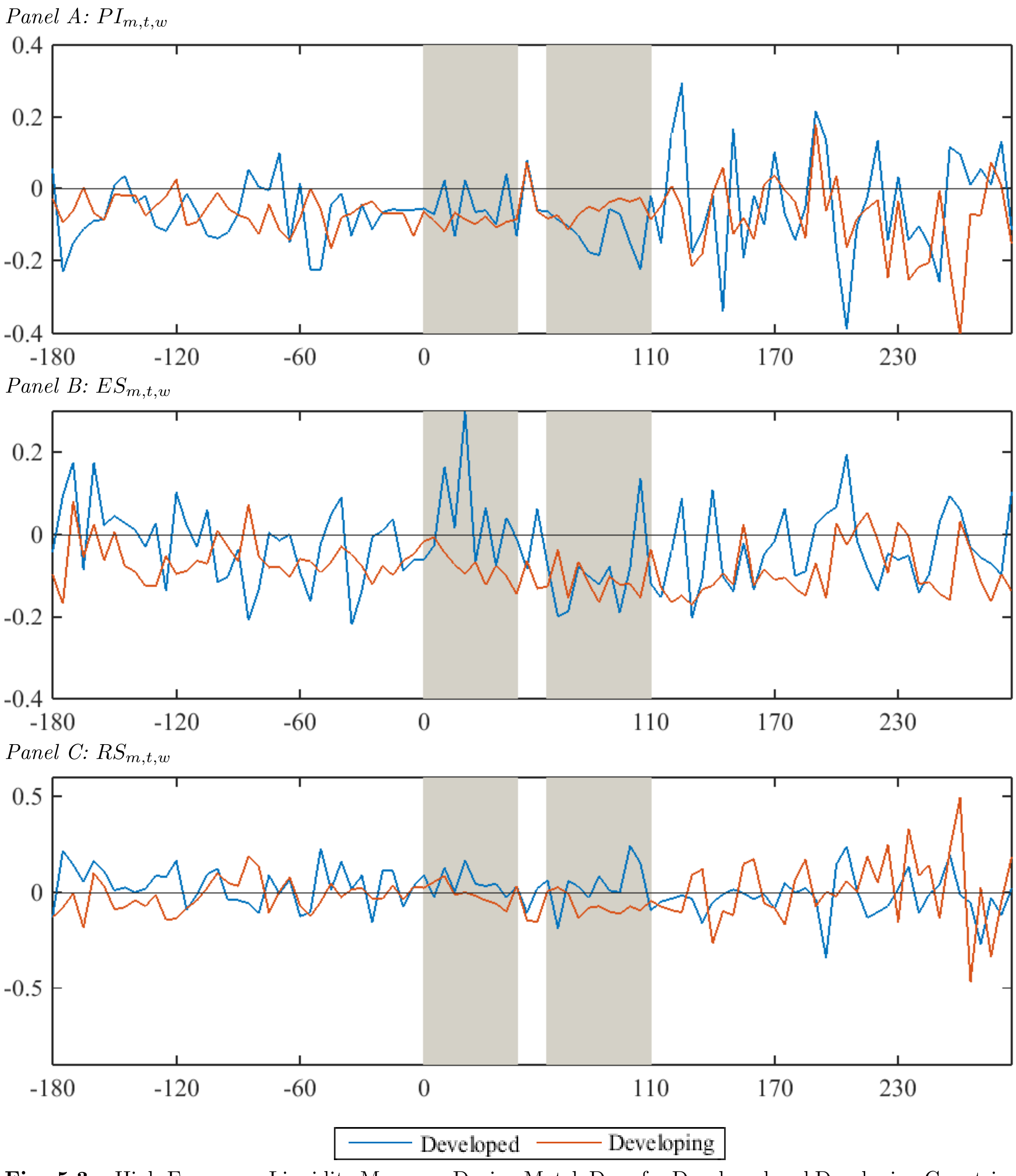


Fig. 5.3. High-Frequency Liquidity Measures During Match Days for Developed and Developing Countries. This figure plots the mean standardised, seasonally detrended high-frequency liquidity variables on match days. The market variables are detrended using the methodology of Gallant et al. (1992) for month-of-the-year, day-of-the-week and time-of-the-day effects. Five-minute price impact for market m at time t and World Cup sub-sample w is denoted by $PI_{m,t,w}$, effective spread by $ES_{m,t,w}$ and realised spread by $RS_{m,t,w}$. The average first-half and second-half time periods are shaded in grey. The x-axis of each plot indicates the number of minutes from kick-off time.

Table 5.2
High Frequency Liquidity Conditions On Match Days

This table reports the estimation results for the following regression:

$$\begin{aligned} DEP_{m,t,w} =& \alpha_0 + \beta_0 GD_{m,t,w} + \sum_{\tau=1}^{6} PRE^{K-\tau}_{m,t,w}\beta_\tau + \beta_7 D^F_{m,t,w} + \beta_8 D^H_{m,t,w} + \beta_9 D^S_{m,t,w} \\ &+ \beta_{10} D^E_{m,t,w} + \sum_{\tau=1}^{6} POST^{F+\tau}_{m,t,w}\beta_{10+\tau} + \epsilon_{m,t,w} \end{aligned} \tag{2.2}$$

where $DEP_{m,t,w}$ is the dependent variable. Each dependent variable is formed by pooling the standardised, detrended market-level observations across World Cup sub-samples. The $GD_{m,t,w}$ indicator variable is a match day indicator variable that takes the value of one if five-minute intra-day time period t coincides with a trading day in which country m is open for trading simultaneous to country m participating in a football match. The $D^F_{m,t,w}$ indicator variable takes the value of one if country m is playing in the first-half of a football match at time t in World Cup sub-sample w. The $D^S_{m,t,w}$ indicator variable is the analogous variable for the second-half of a football match. The $D^H_{m,t,w}$ indicator variable takes the value of one if it is half-time at time t for a match involving country m in World Cup sub-sample w. The $D^E_{m,t,w}$ indicator variable takes the value of one if it is extra-time at time t for a match involving country m in World Cup sub-sample w. The remaining indicator variables capture abnormal trading activity outside of match time on match days. The $PRE^{K-\tau}_{m,t,w}$ indicator variables take the value of one if t is 30τ minutes before a first-half kick-off observation or between 30τ and $30(\tau-1)$ minutes before a first-half kick-off observation, given that country m is participating in the match in World Cup sub-sample w. The $POST^{F+\tau}_{m,t,w}$ indicator variables take the value of one if t is 30τ minutes after a full-time observation or between 30τ and $30(\tau+1)$ minutes after a full-time observation, given that country m is participating in the match in World Cup sub-sample w. Effective Spread for market m at time t during World Cup sub-sample w is denoted by $ES_{m,t,w}$, realised spread by $RS_{m,t,w}$ and five-mute price impact by $PI_{m,t,w}$. The t-statistics are reported in italics. The standard errors are clustered at the country-year level. The 99%, 95%, and 90% confidence levels are indicated by ***, **, and *, respectively. R^2_W is a weighted coefficient of determination that only utilises observations for which at least one non-constant independent variable is nonzero.

$DEP_{m,t,w}$	$ES_{m,t,w}$	$RS_{m,t,w}$	$PI_{m,t,w}$
$GD_{m,t,w}$	0.025	0.012	-0.001
	0.94	*0.80*	*-0.10*
$PRE^{K-6}_{m,t,w}$	0.018	0.006	-0.029*
	0.77	*0.16*	*-1.67*
$PRE^{K-5}_{m,t,w}$	-0.014	-0.017	0.005
	-0.66	*-0.71*	*0.24*
$PRE^{K-4}_{m,t,w}$	0.018	0.040	-0.013
	0.50	*1.20*	*-0.60*
$PRE^{K-3}_{m,t,w}$	-0.003	0.061**	-0.020
	-0.10	*2.08*	*-0.98*
$PRE^{K-2}_{m,t,w}$	-0.013	0.012	-0.036*
	-0.36	*0.37*	*-1.87*
$PRE^{K-1}_{m,t,w}$	-0.037	0.015	-0.022
	-1.34	*0.49*	*-1.07*

Table 5.2
(continued)

$DEP_{m,t,w}$	$ES_{m,t,w}$	$RS_{m,t,w}$	$PI_{m,t,w}$
$D^{F}_{m,t,w}$	-0.016	0.029	-0.045***
	-0.53	*1.07*	*-3.04*
$D^{H}_{m,t,w}$	-0.043	-0.046	0.019
	-1.22	*-1.10*	*0.66*
$D^{S}_{m,t,w}$	-0.080***	-0.027	-0.032*
	-2.82	*-0.90*	*-1.82*
$D^{E}_{m,t,w}$	0.218**	0.349***	-0.172***
	2.47	*16.19*	*-2.66*
$POST^{F+1}_{m,t,w}$	-0.073**	-0.010	-0.015
	-2.23	*-0.34*	*-0.52*
$POST^{F+2}_{m,t,w}$	-0.045	-0.008	-0.025
	-1.41	*-0.18*	*-0.71*
$POST^{F+3}_{m,t,w}$	-0.061*	-0.006	0.036
	-1.78	*-0.17*	*0.77*
$POST^{F+4}_{m,t,w}$	0.047	0.053	-0.043
	0.82	*1.04*	*-1.16*
$POST^{F+5}_{m,t,w}$	-0.006	0.097*	-0.078
	-0.13	*1.78*	*-1.62*
$POST^{F+6}_{m,t,w}$	-0.037	-0.061	0.039
	-0.77	*-1.20*	*0.73*
Observations	239767	239158	256435
R^2_W (%)	0.67	0.35	0.38

Given the surprising results of Figure 5.2, Figure 5.3 and Table 5.2, further research should be dedicated to understanding the dynamics of the high-frequency illiquidity measures. There are a number of potential explanations for the preliminary results presented. Firstly, it could be that the trade and quote data used to calculate that high-frequency illiquidity measures is more prone to reporting errors than the five-minute intra-day data used to calculate the $AMIHUD_{m,t,w}$ and $MAMIHUD_{m,t,w}$ measures. If this is the case, the trade and quote data will need to be filtered in a different manner than the five-minute intra-day data. Thus, future research should apply similar filters and controls to those implemented by Fong et al. (2017), whom use a similar global dataset of trade and quote data from Thomson Reuters Tick History. Second, it could be that some high-frequency illiquidity measures are not appropriate for some countries within the dataset. For example, across the 1998 to 2014 sample period some developing countries are extremely illiquid. With large differences in illiquidity over the country sub-samples, it is not immediately obvious that the five-minute lead applied to the $PI_{i,k}$ and $RS_{i,k}$ measures is equally appropriate for all countries. Finally, given that each high-frequency illiquidity measure captures a different aspect of illiquidity, it could be that some high-frequency illiquidity measures are more relevant to some countries than others. This is because there are large structural differences across the 22 markets in the sample.

5.3. Sports Sentiment and Stock Returns: An Intra-day Study - Limitations and Future Research

The following sub-sections discuss opportunities for further research with respect to Chapter 3 "Sports Sentiment and Stock Returns: An Intra-day Study". Sub-section 5.3.1 outlines plans to improve the intra-day analyses presented in sections 3.4.2 and 3.4.3. Sub-section 5.3.2 discusses options for exploiting cross-sectional variation within the dataset to provide additional evidence of the sentiment effect. Sub-section 5.3.3 outlines plans to examine stock return comovement around shocks to sentiment. Finally, Sub-section 5.3.4 suggests options for identifying retail trades.

5.3.1. A Comparison to Edmans et al. (2007)

The intra-day stock market return results presented in Section 3.4.2 conform with the investor sentiment predictions – Section 3.4.2 finds significant evidence of an intra-day win-effect. Nonetheless, this is in contrast to the daily stock return analysis presented in Section 3.4.1 and Edmans et al. (2007). Section 3.4.1 and Edmans et al. (2007) find a significant inter-day loss effect.[35] Edmans et al. (2007) attribute their asymmetric sentiment effect to a number of factors. First, Edmans et al. (2007) cite the psychology literature: "while an increase in heart attacks, crimes, and suicides is shown to accompany sporting losses, there is no evidence of improvements in mood of a similar magnitude after wins". Second, if football fans suffer from an "allegiance bias" as in Wann, Melnick, Russell, and Pease (2001) and Markman and Hirt (2002), they could place unreasonable expectations on their team winning. Under the reference point argument of Kahneman and Tversky (1979), this could mean that losses have a more pronounced impact on sentiment than wins. This view is supported by Bernile and Lyandres (2011), whom find that "investors overestimate (underestimate) the probability of winning (losing) by nearly 5 percentage points". Third, World Cups are designed with inherent asymmetries. Specifically, when a country losses an Elimination Stage World Cup match, they are immediately removed from the competition; while, the winner merely progresses to the next stage of the competition. This argument is used to motivate the study of Kaplanski and Levy (2010). Thus, it is surprising to find an asymmetric sentiment effect in the opposite direction at the intra-day level. This motivates further research to reconcile the intra-day win-effect with the inter-day loss effect.

Several considerations could explain the contrasting asymmetric sentiment effects. First, the market returns used in Section 3.4.1 and Section 3.4.2 are quite different. Section 3.4.1 makes use of daily total market returns that include all stocks within a market. This data is available from Datastream. In contrast, Section 3.4.2 makes use of intra-day market returns derived from market indices. This data is available from Thomson Reuters Tick History. The sample of market indices included in the intra-day analysis only includes the most prominent market index of each country. Thus, the indices included in the intra-day analysis only include the largest and most

[35] Ashton et al. (2003) only consider a symmetric sentiment effect.

liquid stocks in each country. This distinction could account for the contrasting sentiment results. Thus, further research should seek to expand the cross-section of stocks included in the intra-day stock return analysis. This could be accomplished by selecting market indices with broader market coverage. Expanding the cross-section of stocks could also be achieved by constructing total market indices from the intra-day data available from Thomson Reuters Tick History; however, this option would be rather data intensive. Expanding the cross-section of stock to include less liquid listings is also likely to improve the results presented in Chapter 3. This is because behavioural biases should be stronger for those stocks with sizeable limits to arbitrage (De Long, Shleifer, Summers, and Waldmann (1990); Shleifer and Vishny (1997)). This notion is discussed in more detail in Sub-section 5.3.2

Second, the daily stock return sample and the intra-day stock return sample have contrasting sample periods. The daily stock return sample period is from 1973 to 2014; while, the intra-day stock return sample period is from 1998 to 2014. It could be that since the circulation of the Edmans et al. (2007) study or the preceding working papers, "Football and Stock Returns" by Diego García and Øyvind Norli and "Soccer, Sentiment, and Stocks" by Alex Edmans, the market has sought to gradually correct the documented loss-effect. This notion of a weakening loss-effect is supported by tables 3.8 and 3.9. A weakening loss-effect at the daily level could have been accompanied with a weakening loss-effect at the intra-day level, resulting in the remaining win effect at the intra-day level. Unfortunately, data availability restricts a large longitudinal study of intra-day sentiment effects. Hence, a thorough investigation of this explanation is not feasible in the foreseeable future.

An explanation for the asymmetric sentiment effects can also be developed from arguments presented in a contemporaneous study, Cai et al. (2018). Cai et al. (2018) argue that a significant portion of the loss-effect presented in Edmans et al. (2007) can be attributed to the physiological impacts of sleeplessness that results from investors watching overnight football matches, particularly in the early morning hours. Further, a limited overnight win-effect could be the result of the opposing impacts of positive sentiment and sleeplessness. Given that Section 3.4.2 utilises a sample of football matches that occur during the daylight hours of the participating countries, one should not expect the same asymmetric sentiment effect at the intra-day level. To test this argument, further research should utilise open-to-close returns and close-to-open returns as dependent variables. Under the arguments of Cai et al. (2018), contemporaneous sentiment shocks should have an asymmetric impact on close-to-open returns and a symmetric or more balanced impact on open to close returns.

5.3.2. Cross-sectional Variation and Limits to Arbitrage

Future research should seek to take advantage of cross-sectional variation within the intra-day stock market returns dataset. Ideally, this further research would provide additional evidence of the sentiment effect. For example, behavioural biases should be stronger for those stocks with sizeable limits to arbitrage (De Long et al. (1990); Shleifer and Vishny (1997)). Accordingly, it should be that the intra-day sentiments effects found in Chapter 3 are stronger for those stocks that are more illiquid and volatile. For example, Edmans et al. (2007) and Cai et al. (2018) find that the overnight sports sentiment loss-effect is stronger for small market capitalisation stocks.[36] Thus, further research should involve constructing a small market capitalisation sample and testing whether the sentiment effects are stronger for this sample. Alternatively, I could increase the number of stocks within dataset and perform illiquidity or volatility sorts.

Another opportunity for further research could be to take advantage of country-level variation. For example, further research could perform a sub-sample analysis for those countries with a strong football following. This is an approach taken by Edmans et al. (2007), whom exclude the United States of America and Canada from their analysis. Moreover, Cai et al. (2018) only consider 16 countries with a history of successful participation in the World Cup. Further, it may be possible to exploit variation in economic conditions over the sample. For example, García (2013) shows that sentiment effects can be stronger during recessions. Therefore, future analysis could seek to determine whether the sentiment effects demonstrated in Chapter 3 are related to the economic expansions and contractions of the countries within the sample.

5.3.3. Stock Return Comovement

In addition to the analysis presented in Chapter 3, future research should examine stock return comovement as a variable of interest. This investigation is motivated by previous findings that behavioural biases can induce stock return comovements. For example, Kumar and Lee (2006) use data from a United States discount brokerage firm to demonstrate that retail trades increase stock return comovements. Kumar and Lee (2006) attribute their result to investor sentiment. In another study, Barber, Odean, and Zhu (2009) attribute stock return comovements to behavioural biases such as the extrapolation of past returns, the disposition effect and limited investor attention. Moreover, Pantzalis and Park (2014) find that the comovement of stock returns of firms that are in close geographical proximity is more pronounced in the presence of sports sentiment. Therefore, it could be that stock return comovement increases in reaction to sentiment effects induced by World Cup football matches. If this is the case stock return comovement is likely to be increased following half-time match outcomes and between full-time observations and the close-of-trade. Such an analysis would provide additional evidence of sentiment effects.

5.3.4. Identifying Retail Trades

Finally, more analysis should be conducted to identify the types of traders that are responsible for sentiment effects in financial markets. An appropriate *ex-ante* hypothesis might be that retail traders are more susceptible to sentiment shocks. Without access to exclusive brokerage data or

[36] Surprisingly, Edmans et al. (2007) also find a negative overnight win-effect for their small market capitalisation stock sample.

data with broker IDs, future analysis should make use of retail trade proxies to test whether retail traders are more likely to be influenced by investor sentiment. For example, some researchers have proposed using odd-lot trades to proxy for individuals' trades (Dyl and Maberly (1992); Lakonishok and Maberly (1990); Johnson, Ness, and Ness (2017)).[37] Further, it could be that retail investors are attracted to certain stock properties. For example, Lee, Shleifer, and Thaler (1991) argue that small market capitalisation stocks are predominantly held and traded by individual traders. Kumar and Lee (2006) document that "small firms, lower priced firms, firms with lower institutional ownership, and value (high B/M) firms, all are associated with strong retail concentrations and disproportionately high retail trading activities". If retail investors are more susceptible to sentiment shocks, sentiment effects should be more prevalent among stocks with these characteristics.

[37] O'Hara, Yao, and Ye (2014) argue that odd-lot trades are submitted by sophisticated investors splitting up their trades. The findings of O'Hara et al. (2014) are questioned by Upson and Johnson (2017)

References

Acharya, Viral V., and Lasse Heje Pedersen, 2005, Asset pricing with liquidity risk, *Journal of Financial Economics* 77, 375–410.

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

Akaike, Hirotugu, 1974, A new look at the statistical model identification, *IEEE Transactions on Automatic Control* 19, 716–723.

Alfarano, Simone, Thomas Lux, and Friedrich Wagner, 2008, Time variation of higher moments in a financial market with heterogeneous agents: An analytical approach, *Journal of Economic Dynamics and Control* 32, 101–136.

Alizadeh, Sassan, Michael W. Brandt, and Francis X. Diebold, 2002, Range-based estimation of stochastic volatility models, *The Journal of Finance* 57, 1047–1091.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

Andersen, Torben G., and Tim Bollerslev, 1998, Answering the skeptics: Yes, standard volatility models do provide accurate forecasts, *International Economic Review* 39, 885–905.

Andersen, Torben G., Tim Bollerslev, and Francis X. Diebold, 2007, Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility, *The Review of Economics and Statistics* 89, 701–720.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Heiko Ebens, 2001a, The distribution of realized stock return volatility, *Journal of Financial Economics* 61, 43–76.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys, 2001b, The distribution of realized exchange rate volatility, *Journal of the American Statistical Association* 96, 42–55.

Ashton, J. K., B. Gerrard, and R. Hudson, 2003, Economic impact of national sporting success: evidence from the London Stock Exchange, *Applied Economics Letters* 10, 783–785.

Barber, Brad M., and Terrance Odean, 2008, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *The Review of Financial Studies* 21, 785–818.

Barber, Brad M., Terrance Odean, and Ning Zhu, 2009, Systematic noise, *Journal of Financial Markets* 12, 547–569.

Bekaert, Geert, Campbell R. Harvey, and Christian Lundblad, 2007, Liquidity and expected returns: Lessons from emerging markets, *Review of Financial Studies* 20, 1783–1831.

Berkman, Henk, Paul D. Koch, Laura Tuttle, and Ying Jenny Zhang, 2012, Paying attention: Overnight returns and the hidden cost of buying at the open, *Journal of Financial and Quantitative Analysis* 47, 715–741.

Bernile, Gennaro, and Evgeny Lyandres, 2011, Understanding investor sentiment: The case of soccer, *Financial Management* 40, 357–380.

Beveridge, Stephen, and Charles R. Nelson, 1981, A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle', *Journal of Monetary Economics* 7, 151–174.

Black, Fisher S., 1976, Studies of stock market volatility changes, *Proceedings of the American Statistical Association* Business and Economic Statistics Section, 177–181.

Bouchaud, Jean-Philippe, J. Doyne Farmer, and Fabrizio Lillo, 2009, How markets slowly digest changes in supply and demand, *Handbook of financial markets: Dynamics and evolution* 57–160.

Bouchaud, Jean-Philippe, Yuval Gefen, Marc Potters, and Matthieu Wyart, 2004, Fluctuations and response in financial markets: The subtle nature of 'random' price changes, *Quantitative Finance* 4, 176–190.

Brailsford, Timothy J., 1996, The empirical relationship between trading volume, returns and volatility, *Accounting & Finance* 36, 89–111.

Brunnermeier, Markus K., and Lasse Heje Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201–2238.

Cai, Jinghan, Manyi Fan, Chiu Yu Ko, Marco Richione, and Natalie Russo, 2018, Physiology, psychology, and stock market performance: Evidence from sleeplessness and distraction in the world cup, Working Paper, University of Scranton.

Campbell, John Y., and Ludger Hentschel, 1992, No news is good news: An asymmetric model of changing volatility in stock returns, *Journal of Financial Economics* 31, 281–318.

Chae, Joon, 2005, Trading volume, information asymmetry, and timing information, *The Journal of Finance* 60, 413–442.

Chan, Louis K. C., and Josef Lakonishok, 1993, Institutional trades and intraday stock price behavior, *Journal of Financial Economics* 33, 173–199.

Chan, Louis K. C., and Josef Lakonishok, 1995, The behavior of stock prices around institutional trades, *The Journal of Finance* 50, 1147–1174.

Chang, Shao-Chi, Sheng-Syan Chen, Robin K. Chou, and Yueh-Hsiang Lin, 2012, Local sports sentiment and returns of locally headquartered stocks: A firm-level analysis, *Journal of Empirical Finance* 19, 309 – 318.

Chiarella, Carl, Giulia Iori, and Josep Perelló, 2009, The impact of heterogeneous trading rules on the limit order book and order flows, *Journal of Economic Dynamics and Control* 33, 525–537.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2000, Commonality in liquidity, *Journal of Financial Economics* 56, 3–28.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2001a, Market liquidity and trading activity, *The Journal of Finance* 56, 501–530.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2002, Order imbalance, liquidity, and market returns, *Journal of Financial Economics* 65, 111–130.

Chordia, Tarun, Asani Sarkar, and Avanidhar Subrahmanyam, 2005, An empirical analysis of stock and bond market liquidity, *Review of Financial Studies* 18, 85–129.

Chordia, Tarun, Asani Sarkar, and Avanidhar Subrahmanyam, 2011, Liquidity dynamics and cross-autocorrelations, *Journal of Financial and Quantitative Analysis* 46, 709–736.

Chordia, Tarun, Avanidhar Subrahmanyam, and V. Ravi Anshuman, 2001b, Trading activity and expected stock returns, *Journal of Financial Economics* 59, 3–32.

Christie, Andrew A., 1982, The stochastic behavior of common stock variances: Value, leverage and interest rate effects, *Journal of Financial Economics* 10, 407–432.

Corsi, Fulvio, 2009, A simple approximate long-memory model of realized volatility, *Journal of Financial Econometrics* 7, 174–196.

Corsi, Fulvio, and Roberto Renó, 2012, Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling, *Journal of Business and Economic Statistics* 30, 368–380.

Covitz, Dan, and Chris Downing, 2007, Liquidity or credit risk? The determinants of very short-term corporate yield spreads, *The Journal of Finance* 62, 2303–2328.

Damodaran, Aswath, 1989, The weekend effect in information releases: A study of earnings and dividend announcements, *The Review of Financial Studies* 2, 607–623.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703–738.

Dellavigna, Stefano, and Joshua M. Pollet, 2009, Investor inattention and Friday earnings announcements, *The Journal of Finance* 64, 709–749.

Deo, Rohit, Clifford Hurvich, and Yi Lu, 2006, Forecasting realized volatility using a long-memory stochastic volatility model: Estimation, prediction and seasonal adjustment, *Journal of Econometrics* 131, 29–58.

Domowitz, Ian, Jack Glen, and Ananth Madhavan, 2001, Liquidity, volatility and equity trading costs across countries and over time, *International Finance* 4, 221–255.

Drienko, Jozef, Tom Smith, and Anna von Reibnitz, 2018, A review of the return–illiquidity relationship, *Critical Finance Review.* Forthcoming.

Dyl, Edward A., and Edwin D. Maberly, 1992, Odd-lot transactions around the turn of the year and the January effect, *Journal of Financial and Quantitative Analysis* 27, 591–604.

Edmans, Alex, Diego García, and Øyvind Norli, 2007, Sports sentiment and stock returns, *The Journal of Finance* 62, 1967–1998.

Ehrmann, Michael, and David-Jan Jansen, 2016, It hurts (stock prices) when your team is about to lose a soccer match, *Review of Finance* 20, 1215–1233.

Ehrmann, Michael, and David-Jan Jansen, 2017, The pitch rather than the pit: Investor inattention, trading activity, and FIFA World Cup matches, *Journal of Money, Credit and Banking* 49, 807–821.

Eiichiro Kazumori, Raj Sharman Fumiko Takeda, Fei Fang, and Hong Yu, 2018, Asset pricing with liquidity risk: A replication and out-of-sample tests with the recent US and the Japanese market data, Working Paper, University of Massachusetts.

Elo, Arpad Emmerich, 1978, *The rating of chessplayers, past and present* (Arco Pubishing).

Evans, Twm, and David G. McMillan, 2009, Asymmetric return patterns: Evidence from 33 international stock market indices, *Applied Economics Letters* 16, 775–779.

Fong, Kingsley Y L, Craig W Holden, and Charles A Trzcinka, 2017, What are the best liquidity proxies for global research?, *Review of Finance* 21, 1355–1401.

Foster, F. Douglas, and S. Viswanathan, 1990, A theory of the interday variations in volume, variance, and trading costs in securities markets, *The Review of Financial Studies* 3, 593–624.

Foster, F. Douglas, and S. Viswanathan, 1993, Variations in trading volume, return volatility, and trading costs: Evidence on recent price formation models, *The Journal of Finance* 48, 187–211.

French, Kenneth R., G. William Schwert, and Robert F. Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 19, 3–29.

Fung, Hung-Gay, and Gary A. Patterson, 1999, The dynamic relationship of volatility, volume, and market depth in currency futures markets, *Journal of International Financial Markets, Institutions and Money* 9, 33–59.

Gallant, A. Ronald, Peter E. Rossi, and George Tauchen, 1992, Stock prices and volume, *The Review of Financial Studies* 5, 199–242.

García, Diego, 2013, Sentiment during recessions, *The Journal of Finance* 68, 1267–1300.

Gerig, Austin N., 2007, *A Theory of Market Impact: How Order Flow Affects Stock Price*, Ph.D. thesis, University of Illinois.

Gillemot, László, J. Doyne Farmer, and Fabrizio Lillo, 2006, There's more to volatility than volume, *Quantitative Finance* 6, 371–384.

Goyenko, Ruslan Y., Craig W. Holden, and Charles A. Trzcinka, 2009, Do liquidity measures measure liquidity?, *Journal of Financial Economics* 92, 153–181.

Hameed, Allaudeen, Wenjin Kang, and S. Viswanathan, 2010, Stock market declines and liquidity, *The Journal of Finance* 65, 257–293.

Hansen, Bruce E., 1997, Approximate asymptotic p values for structuraS-change tests, *Journal of Business and Economic Statistics* 15, 60–67.

Harris, Larry, and Andrea Amato, 2018, Illiquidity and stock returns: Cross-section and time-series effects: A replication, *Critical Finance Review.* Forthcoming.

Hasbrouck, Joel, 1991, The summary informativeness of stock trades: An econometric analysis, *The Review of Financial Studies* 4, 571–595.

Hasbrouck, Joel, and Duane J. Seppi, 2001, Common factors in prices, order flows, and liquidity, *Journal of Financial Economics* 59, 383–411.

Hirshleifer, David, Sonya Seongyeon Lim, and Siew Hong Teoh, 2009, Driven to distraction: Extraneous events and underreaction to earnings news, *The Journal of Finance* 64, 2289–2325.

Holden, Craig W., and Jayoung Nam, 2018, Do the LCAPM predictions hold? Replication and extension evidence, Working Paper, Indiana University.

Huberman, Gur, and Dominika Halka, 2001, Systematic liquidity, *Journal of Financial Research* 24, 161–178.

Hurst, Harold Edwin, 1951, Long-term storage capacity of reservoirs, *American Society of Civil Engineers* 116, 770–799.

Johnson, Hardy, Bonnie F. Van Ness, and Robert A. Van Ness, 2017, Are all odd-lots the same? odd-lot transactions by order submission and trader type, *Journal of Banking & Finance* 79, 1–11.

Kahneman, Daniel, and Amos Tversky, 1979, Prospect theory: An analysis of decision under risk, *Econometrica* 47, 263–292.

Kamstra, Mark J., Lisa A. Kramer, and Maurice D. Levi, 2000, Losing sleep at the market: The daylight saving anomaly, *American Economic Review* 90, 1005–1011.

Kamstra, Mark J., Lisa A. Kramer, and Maurice D. Levi, 2002, Losing sleep at the market: The daylight saving anomaly: Reply, *American Economic Review* 92, 1257–1263.

Kaplanski, Guy, and Haim Levy, 2010, Exploitable predictable irrationality: The FIFA World Cup effect on the u.s. stock market, *Journal of Financial and Quantitative Analysis* 45, 535–553.

Kaplanski, Guy, and Haim Levy, 2014, Sentiment, irrationality and market efficiency: The case of the 2010 FIFA World Cup, *Journal of Behavioral and Experimental Economics* 49, 35 – 43.

Kaplanski, Guy, Haim Levy, Chris Veld, and Yulia Veld-Merkoulova, 2015, Do happy people make optimistic investors?, *Journal of Financial and Quantitative Analysis* 50, 145–168.

Kumar, Alok, and Charles M.C. Lee, 2006, Retail investor sentiment and return comovements, *The Journal of Finance* 61, 2451–2486.

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

Kyle, Albert S., Anna A. Obizhaeva, and Yajun Wang, 2014, Smooth trading with overconfidence and market power, Working Paper, University of Maryland.

Kyle, Albert S., and Wei Xiong, 2001, Contagion as a wealth effect, *The Journal of Finance* 56, 1401–1440.

Lakonishok, Josef, and Edwin Maberly, 1990, The weekend effect: Trading patterns of individual and institutional investors, *The Journal of Finance* 45, 231–243.

LeBaron, Blake, and Ryuichi Yamamoto, 2007, Long-memory in an order-driven market, *Physica A: Statistical Mechanics and its Applications* 383, 85–89.

Lee, Charles M. C., and Mark J. Ready, 1991, Inferring trade direction from intraday data, *The Journal of Finance* 46, 733–746.

Lee, Charles M. C., Andrei Shleifer, and Richard H. Thaler, 1991, Investor sentiment and the closed-end fund puzzle, *The Journal of Finance* 46, 75–109.

Levine, Ross, and Sara Zervos, 1998, Capital control liberalization and stock market development, *World Development* 26, 1169–1183.

Lillo, Fabrizio, Szabolcs Mike, and J. Doyne Farmer, 2005, Theory for long memory in supply and demand, *Physical Review E* 71, 066122.

Liu, Wai-Man, 2009, Monitoring and limit order submission risks, *Journal of Financial Markets* 12, 107–141.

Lo, Andrew W., 1991, Long-term memory in stock market prices, *Econometrica* 59, 1279–1313.

Lou, Xiaoxia, and Tao Shu, 2017, Price impact or trading volume: Why is the Amihud (2002) illiquidity measure priced?, *Review of Financial Studies,* forthcoming.

Mancini-Griffoli, Tommaso, and Angelo Ranaldo, 2011, Limits to arbitrage during the crisis: funding liquidity constraints and covered interest parity, Working Paper, University of St. Gallen.

Markman, Keith, and Edward Hirt, 2002, Social prediction and the "allegiance bias", *Social Cognition* 20, 58–86.

Martens, Martin, Dick Van Dijk, and Michiel De Pooter, 2009, Forecasting S&P 500 volatility: Long memory, level shifts, leverage effects, day-of-the-week seasonality, and macroeconomic announcements, *International Journal of Forecasting* 25, 282–303.

McAleer, Michael, and Marcelo C. Medeiros, 2008, A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries, *Journal of Econometrics* 147, 104–119.

Mike, Szabolcs, and J. Doyne Farmer, 2008, An empirical behavioral model of liquidity and volatility, *Journal of Economic Dynamics and Control* 32, 200–234.

Mishra, Vinod, and Russell Smyth, 2010, An examination of the impact of India's performance in one-day cricket internationals on the Indian stock market, *Pacific-Basin Finance Journal* 18, 319–334.

Mitchell, Mark, Lasse Heje Pedersen, and Todd Pulvino, 2007, Slow moving capital, *American Economic Review* 97, 215–220.

Müller, Ulrich A., Michel M. Dacorogna, Rakhal D. Davé, Olivier V. Pictet, Richard B. Olsen Olsen, and Robert Ward, 1993, Fractals and intrinsic time - A challenge to econometricians, in *39th International AEA Conference on Real Time Econometrics, 14–15 October 1993, Luxembourg*.

Nagel, Stefan, 2012, Evaporating liquidity, *Review of Financial Studies* 25, 2005–2039.

Nam, Kiseok, Kenneth M. Washer, and Quentin C. Chu, 2005, Asymmetric return dynamics and technical trading strategies, *Journal of Banking and Finance* 29, 391–418.

O'Hara, Maureen, Chen Yao, and Mao Ye, 2014, What's not there: Odd lots and market data, *The Journal of Finance* 69, 2199–2236.

Pantzalis, Christos, and Jung Chul Park, 2014, Exuberance out of left field: Do sports results cause investors to take their eyes off the ball?, *Journal of Economic Behavior & Organization* 107, Part B, 760–780, Empirical Behavioral Finance.

Parkinson, Michael, 1980, The extreme value method for estimating the variance of the rate of return, *The Journal of Business* 53, 61–65.

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

Patell, James M., and Mark A. Wolfson, 1982, Good news, bad news, and the intraday timing of corporate disclosures, *The Accounting Review* 57, 509–527.

Patton, Andrew J., 2011, Volatility forecast comparison using imperfect volatility proxies, *Journal of Econometrics* 160, 246–256.

Patton, Andrew J., and Kevin Sheppard, 2013, Good volatility, bad volatility: Signed jumps and the persistence of volatility, Working Paper, University of Oxford.

Penman, Stephen H., 1987, The distribution of earnings news over time and seasonalities in aggregate stock returns, *Journal of Financial Economics* 18, 199–228.

Pinegar, J. Michael, 2002, Losing sleep at the market: Comment, *American Economic Review* 92, 1251–1256.

Plerou, Vasiliki, Parameswaran Gopikrishnan, and H. Eugene Stanley, 2005, Quantifying fluctuations in market liquidity: Analysis of the bid-ask spread, *Physical Review E* 71, 046131.

Rhee, S. Ghon, and Jianxin Wang, 2009, Foreign institutional ownership and stock market liquidity: Evidence from Indonesia, *Journal of Banking and Finance* 33, 1312–1324.

Scalia, Antonio, 1998, Periodic information asymmetry and intraday market behaviour: An empirical analysis, *Review of Finance* 1, 307–305.

Schwarz, Gideon, 1978, Estimating the dimension of a model, *The Annals of Statistics* 6, 461–464.

Seasholes, Mark S., and Guojun Wu, 2007, Predictable behavior, profits, and attention, *Journal of Empirical Finance* 14, 590 – 610.

Sévi, Benoît, 2014, Forecasting the volatility of crude oil futures using intraday data, *European Journal of Operational Research* 235, 643–659.

Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *The Journal of Finance* 52, 35–55.

Stoll, Hans R., and Robert E. Whaley, 1983, Transaction costs and the small firm effect, *Journal of Financial Economics* 12, 57–79.

Upson, James, and Hardy Johnson, 2017, Are odd-lot orders informed?, *Financial Review* 52, 37–67.

Wang, George H. K., and Jot Yau, 2000, Trading volume, bid-ask spread, and price volatility in futures markets, *Journal of Futures Markets* 20, 943–970.

Wang, Jessica, and Raphael N. Markellos, 2015, Is there an Olympic gold medal rush in the stock market, Working Paper, University of East Anglia.

Wang, Jianxin, 1999, Asymmetric information and the bid-ask spread: An empirical comparison between automated order execution and open outcry auction, *Journal of International Financial Markets, Institutions and Money* 9, 115–128.

Wang, Jianxin, 2011, Forecasting volatility in Asian stock markets: Contributions of local, regional, and global factors, *Asian Development Review* 28, 32–57.

Wang, Jianxin, 2013, Liquidity commonality among Asian equity markets, *Pacific-Basin Finance Journal* 21, 1209–1231.

Wang, Jianxin, and Minxian Yang, 2009, Asymmetric volatility in the foreign exchange markets, *Journal of International Financial Markets, Institutions and Money* 19, 597–615.

Wang, Jianxin, and Minxian Yang, 2017, Conditional volatility persistence, Working paper, University of Technology Sydney.

Wann, Daniel L., Merrill J. Melnick, Gordon W. Russell, and Dale G. Pease, 2001, *Sport fans: The psychology and social impact of spectators*, volume 1 (Routledge New York).

Weber, Philipp, and Bernd Rosenow, 2006, Large stock price changes: Volume or liquidity?, *Quantitative Finance* 6, 7–14.

Wyart, Matthieu, and Jean-Philippe Bouchaud, 2007, Self-referential behaviour, overreaction and conventions in financial markets, *Journal of Economic Behavior and Organization* 63, 1–24.

Xiong, Wei, 2001, Convergence trading with wealth effects: An amplification mechanism in financial markets, *Journal of Financial Economics* 62, 247–292.

Yamamoto, Ryuichi, 2011, Order aggressiveness, pre-trade transparency, and long memory in an order-driven market, *Journal of Economic Dynamics and Control* 35, 1938–1963.