Price Discovery and Information Asymmetry in Equity and Commodity Futures Options Markets

Marc J.M. Bohmann

A dissertation submitted for the degree of

Doctor of Philosophy

Discipline of Finance

University of Technology Sydney

January 1st, 2020

Fake it till you make it.

Acknowledgments

I am deeply thankful to Professor Dave Michayluk for his mentorship, supervision, comments and overall assistance during the process of researching and writing this thesis. I am also deeply indebted to Dr Vinay Patel for his continuous guidance and help with writing this work. I thank the University of Technology Sydney (UTS) Business School and the Capital Markets Cooperative Research Centre (CMCRC) for their technological and financial support, provision of data and for, more generally, making this work possible.

I would like to thank Bart Frijns, Vito Mollica, Dmitriy Muravyev, Christina Nikitopoulos, Talis Putniņš, Tom Smith, Gerard Van de Venter, Kathy Walsh and Terry Walter for their helpful comments and assistance. In addition, I want to thank my fellow Ph.D. students at UTS and the CMCRC, especially Nihad Aliyev, Edward Curran, Aldo Dalla Costa, Riccardo de Blassis, Stephen Friel, Junqing Kang, Jonathan Karlsen, Martha Khomyn, Juliane Krug and Florian Schroeder.

I appreciate the helpful feedback I received from the following conferences: CMCRC Research Presentations 2016–2018 (Sydney), Singapore Economic Review Conference 2017, ACFR Derivative Markets Conference 2018 (Auckland), FMA Annual Meeting 2018 (San Diego), Vietnam Conference in Finance 2018 (Ho Chi Minh City) and International Accounting & Finance Doctoral Symposium 2019 (Milan). Additionally, I feel grateful and honoured that Chapter 3 is published in the Journal of Derivatives, Chapter 4 is published in the Journal of Futures Markets and Chapter 5 is published in the Pacific-Basin Finance Journal.

Also, I would like to thank my mother (Dorit), father (Andreas), sister (Laura) and grandparents (Josef and Irmgard Kehmer and Ursula Haas) in Germany for their endless support, love and trust in me. Finally, I thank all my friends in Germany and around the world without whom I would not have made to where I am now, especially the entire Immer Lustig-group and OG-crew!

Capstone Editing provided copyediting and proofreading services, according to the guidelines laid out in the university-endorsed national 'Guidelines for Editing Research Theses'.

Certificate of Authorship

I certify that the work in this dissertation has not previously been submitted for a degree, nor has it been submitted as part of the requirements for a degree except as fully acknowledged within the text.

I also certify that the dissertation has been written by me. Any help that I have received in my research work and the preparation of the dissertation has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This research is supported by the Australian Government Research Training Program.

Signature

Production Note:
Signature removed prior to publication.

Marc Bohmann, 1 January 2020

Abstract

This dissertation contributes to the existing literature by examining trading behaviour around security-level and market-wide events. The research focuses on equity and futures options and continues by providing insights into the price discovery process of futures and options in commodity markets.

The first essay examines informed options trading around a sample of 352 Food and Drug Administration announcements from 166 United States (US)-listed firms between 1996 and 2016. Using implied volatility (IV) spreads and options trading volume as proxies for informed trading, it is found that informed traders are aware of the timing of the upcoming announcement at least five days in advance and at least some informed traders have knowledge of the finer details that affect the price impact of the announcement. These findings have implications for regulators, investors and relevant firms.

The second essay analyses the behaviour of United States (US) commodity futures and options IV-based measures as proxies for information leakage around macro-economic and commodity-specific news announcements between 2007 and 2017. In the three days preceding news releases, abnormal changes in the levels of futures options IV spreads and skew were reported. In addition, a statistically significant relationship between announcement date returns and abnormal changes in pre-announcement IV spreads and skew were reported. Universally, the findings indicate that at least some investors are informed about the price impact of the upcoming news announcements in seven commodity markets.

The third essay investigates the extent of the importance of commodity futures or options markets in the price discovery process in the six most-actively traded markets: crude oil, natural gas, gold, silver, corn and soybeans. Using new information and leadership techniques, new evidence has reported that, in recent times, both markets make a meaningful contribution to price discovery. However, on average, options lead futures in reflecting new information for most of these commodities. In addition, it was found that increased speculation—rather than hedging activity—in commodity derivatives is a key determinant of price discovery in the options markets.

The fourth essay sheds light on high versus low-frequency (LF) liquidity measures in times of information asymmetry. Market microstructure data availability has significantly improved and it is now possible to estimate liquidity measures at the nanosecond level. However, this level of data are not available in all markets and time periods and there is a significant cost and computational burden of high-frequency (HF) data. Goyenko et al. (2009) and Fong et al. (2017) show that various LF liquidity measures can proxy for HF benchmarks and show that the results are robust across countries and time. However, liquidity measures do not always behave in the expected fashion during periods of information asymmetry (Collin-Dufresne & Fos, 2015). Drawing from Ball and Brown (1968), an event study methodology is used to investigate whether the LF measures of liquidity can proxy for HF measures around earnings announcements (i.e., periods of information asymmetry). It was found that the closing-percent-quoted-spread is the best proxy for the percent–cost HF benchmarks. In contrast, using cross-sectional, portfolio and individual time-series correlations, the most consistent LF proxies are the high-low impact and closing-percent-quoted-spread impact. However, the performance of these proxies weakens in the pre- and post-announcement periods around the earnings announcement.

Contents

List of Figures

List of Tables

List of Abbreviations

ARET	average abnormal return
ASX	Australian Securities Exchange
ATM	at-the-money
CBOE	Chicago Board of Options Exchange
CBOT	Chicago Board of Trade
CFTC	Commodity Futures Trading Commission
CMCRC	Capital Markets Cooperative Research Centre
CME	Chicago Mercantile Exchange
COMEX	Commodity Exchange
COT	Commitment of Traders
CPI	Consumer Price Index
CPQS	closing-percent-quoted spread
CPQSI	closing-percent-quoted-spread-impact
CS	component share
DITM	deep in the money
DOTM	deep out of the money
FDA	Food and Drug Administration
GDP	gross domestic product
GFC	global financial crisis
HF	high-frequency
HL	high-low
HLI	high-low-impact
ILS	information leadership share
IP	industrial production
IS	information share
IV	implied volatility
LF	low-frequency
MAPP	Manual of Policies and Procedures
NDA	new drug approvals
NYMEX	New York Mercantile Exchange
OTM	out-of-the-money

OTC	over the counter
QTT	quote-to-trade ratio
RFQ	request for quote
RMSE	root-mean-squared-error
SEC	Securities and Exchange Commission
SIC	Standard Industrial Classification
SVOL	stock volume
TRTH	Thomson Reuters Tick History
US	United States
UTS	University of Technology Sydney
VECM	vector error correction model
WASDE	World Agricultural Supply and Demand Estimates Report

Chapter 1: Introduction

1.1 Overview

In the field of economics, demand and supply are the primary forces that establish prices, which reflect the revelation of information to the public. Finance branched out from economics by demonstrating that information is reflected in security prices, which augments demand and supply. One example of the force of information is in work by Ball and Brown (1968), who linked accounting information to share prices. However, Hayek (1945) argued that the full set of price information is not known to every individual investor. Therefore, information asymmetry arises in the markets.

There are many streams of literature that relate information to pricing. The concept of diversification was applied, which led to the capital asset pricing model and the concept of beta—non-diversifiable risk—that must be compensated to hold a risky asset. One branch of this literature examines the process by which information enters prices; the so-called price discovery process. Price discovery is defined as the process of efficient and timely incorporation of information revealed by the trading activity of investors (Lehmann, 2002). Price discovery is the focus of this research, which can be examined in all asset classes, including commodities (Figuerola-Ferretti & Gonzalo, 2010) and derivatives (Sarwar, 2005; Yang & Bessler, 2004).

There are some impediments to converting information into prices. In efficient markets, all information is available to all participants, which directly affects price formation (Malkiel & Fama, 1970). However, this 'perfect world' does not exist, given the amount of early-informed parties involved and the nature of information flow. O'Hara (2003, p. 1351) addressed this issue by stating that 'full information levels is not credible—new information arrives, old information becomes stale and even informed traders may face risks that their information is obsolete'. Early work by Grossman and Stiglitz (1980) and French and Roll (1986) argued that the traditional view of an efficient market does not hold, due to informational asymmetry, mispricing and imprecision of the asset valuations of informed traders. Additionally, Cohen et al. (1983) found that trading frictions, such as trading costs, affect security prices in a way in which returns are asynchronously generated. Similarly, in models by Kyle (1985) and Easley and O'Hara (1987), trade size and quantity inflicted by informed traders affect security prices due to inventory balancing of market makers away from efficiency.

Under current law in most Western countries, insiders with private knowledge about fundamental value, causing information asymmetry, are prohibited from trading on this information prior to release to the public. Therefore, information assimilation into prices is delayed. The general informativeness, measured by return variation, increases significantly (Fernandes & Ferreira, 2008; Kusnadi, 2015). However, in any market, investors with private information about the fundamental value of an asset can make use of it to profit substantially in capital markets (Brunnermeier, 2005). Early trading by insiders through so-called information leakage is one part of the analyses to measure price efficiency and transparency on a market microstructure level. Especially after the enormous growth of capital market volumes and the increase in trading frequency, information and prices are closely related (Adams & Kartsakli, 2017; Tang & Xiong, 2012). Thus, the analyses are conducted, not only daily, but also on an intraday, trade and quote level. This thesis will illuminate a variety of trading characteristics such as returns, volatility and liquidity.

The expansion of related financial assets, such as derivatives and the increase in trading venues, means that information about one asset has many different venues through which it can enter the market. This increase in venues allows an in-depth consideration of information assimilation through the exploitation of arbitrage opportunities across markets and the utilisation of market specifications such as leverage and anonymity. In addition, this thesis is interested in the preferred trading venues used in today's high-frequency (HF) markets. Therefore, it is estimated that when information is first transformed into prices, price discovery occurs. In this regard, investors face the decision of whether to trade in the underlying stock or futures market or the corresponding options market. The derivatives markets are the primary focus of this research, due to their distinct pricing, leverage and strategic characteristics (Black, 1975, Easley, O'Hara & Srinivas, 1998). In particular, this research fills a gap in the literature by focusing on the commodity derivative markets, reflecting the enormous growth in trading activity and its increasing importance (Basu & Gavin, 2010).

This chapter will introduce the relevant markets and explain the principles that led to an understanding of the process of how information is obtained and incorporated into prices.

The current methods used to detect the preferred trading venues of traders is described. The way in which informational efficiency, leakage and transformation are intertwined is demonstrated. Finally, the overarching aspect of information asymmetry and liquidity are linked with the price discovery process and the concept of information leakage.

1.2 Market Microstructure

1.2.1 Differences Between Equity and Futures Options

Although options on equities and futures can serve the same purpose, the two vehicles have different structures. First, unlike equity options, futures options and the underlying asset are traded on the same exchange. This allows a perfect fit of trading hours and settlement. Second, the pricing behaviour of futures options differs due to the bi-directional risk faced by commodity derivative traders. Although stocks have only downside risk, the risk reflected by the IV in futures options is not linear, but parabolic. Third, the structure of futures options varies by underlying security. In contrast, equity options are created with a fixed basis of 100 shares as underlying on the major US exchanges.

The major difference between equity and futures options on a microstructural level is the quoting technique. The quotes of equity options prices are directly linked to the underlying stock and are updated automatically. In contrast, futures options quotes on the Chicago Mercantile Exchange (CME) are created using a request for quote (RFQ), which is comparable to the OTC method. Traders can access bids, request prices and submit requests for quotes through their brokerage platform.[1] Section 4.2 will highlight this difference in greater detail due to the importance of price discovery between futures and corresponding options.

[1] More information can be found on the exchange website: https://www.cmegroup.com/education/request-for-quote.html.

1.2.2 Volume Growth and Financialisation

There has been significant growth in commodity derivative trading activity due to the development of electronic markets (Adams & Kartsakli, 2017; Tang & Xiong, 2012). The Futures Industry Association reported that global commodity futures volume increased sixfold from one to six billion contracts between 2007 and 2016. During this period, the volume of commodity futures options also increased from 100 million to 700 million contracts, while the volume of equity options decreased (Acworth, 2017; Simon, 2014). Futures option volume represents approximately 10 to 20 percent of the total commodity derivatives volume—a substantial fraction. However, relatively little is known about how information is incorporated into futures options prices. Boyd and Locke (2014) are an exception, concluding that, during the period of their study (2005–2007), most information is first reflected in natural gas futures, rather than options contracts.

Chari and Christiano (2017) described financialisation as a development in which derivative trading grows faster than the underlying commodity spot market. In more recent times, and due to advances in technology, the financialisation of commodities and the development of electronic trading platforms has been observed, which have significantly improved the ease of trading commodities. These two factors have led to growth in trading volumes in futures and their corresponding options contracts (CME Group Inc., 2016). For example, the TABB Group reported that the total futures options volume has grown almost sevenfold from 1999 to more than USD700 million in contracts in 2014 (Simon, 2014).

1.2.3 Option Pricing and Implied Volatility

On most trading platforms, the pricing of an option is best represented by the Black-Scholes-Merton model from 1973. This model is used to calculate the theoretical value of European style options. Various deviations exist that account for American style options and other input parameters.[2] The original model uses the underlying price, dividends, options strike price, option time to expiration, expected volatility and interest

[2] European style and American style options only differ in the ability to exercise the respective options contract. European options are option contracts that can only be exercised at the expiration date. American options can be exercised at any time before the expiration date.

rate. According to the model for stock S_t at time t, the price of a European style call (C) and put (P) option can be expressed as:

$$C(S_t, t) = S_t N(d_1) - Ke^{-rT} N(d_2) \tag{1.1}$$

$$P(S_t, t) = Ke^{-rT} N(-d_2) - S_t N(-d_1) \tag{1.2}$$

in which

$$d_1 = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}} \tag{1.3}$$

$$d_2 = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T} \tag{1.4}$$

N is the cumulative distribution function of the standard normal distribution, T is the time to maturity in years, K is the strike price of the option, r is the annual continuous compounding risk-free rate and σ is the return volatility of the underlying asset.

The variable of interest in the Black-Scholes-Merton model is the expected, or implied, volatility (σ). The term cannot be directly plugged into the formula and can only be estimated from the underlying asset price behaviour. In practice, traders use the observed option prices in the markets. Therefore, IV is calculated using an iterative process or more advanced techniques such as the Cox-Ross-Rubinstein (1979) or Newton-Rhapson (Manaster & Koehler, 1982) method. IV reflects the 'market's opinion about the volatility of a stock'. IVs are forward-looking and are used more frequently in option quotes due to their strong effect on the prices of the options and due to the smaller variability than options prices.

1.3 Informed Trading and Information Leakage

Informed trading is the act on which a trader executes trades based on private information. The information could have been obtained legally or illegally. Informed trading, or insider trading, becomes illegal when information about the security, such as a stock, is material and non-public. Regulators, such as the US Securities and Exchange Commission (SEC) or the Commodity Futures Trading Commission (CFTC), enforce the corresponding laws and regulations, such as SEC Rules 10b-5 and 10b-5-1, which define insider trading. The process of information acquisition can include information leakage, which describes the presumably unintentional revelation of information to an unauthorised party. This party can then trade on leaked information for a profit based on an unfair advantage over outside investors (Bhattacharya & Daouk, 2002). However, the extent, benefits and detriments of insider trading that result from information leakage have been debated for decades, with researchers arguing for both sides of the discussion (Bainbridge, 2002).

A recent stream of literature provides support for options as an important venue for informed trading. This evidence begins with a sequential trading model developed by Easley et al. (1998); that is, informed traders trade in options for sufficient levels of liquidity and leverage of options contracts. Second, several studies found that future stock returns can be predicted from options volume or order imbalance (Ge, Lin & Pearson, 2016; Johnson & So, 2012; Pan & Poteshman, 2006). Third, numerous studies reported evidence of abnormal options returns, volume or IV prior to a variety of corporate news announcements (Chan, Ge & Lin, 2015; Christophe, Ferri & Hsieh, 2010; Hao, 2016; Roll, Schwartz & Subrahmanyam, 2010). More recently, using an adaptation of information shares, Patel, Putniņš, Michayluk and Foley (2018) found that approximately one-third of price discovery occurs in options relative to stocks. This estimate is up to five times larger than what has been reported in prior studies, but is consistent with the theory about the preference for leveraged financial instruments.[3] Additionally, many SEC insider trading prosecutions revealed that options are a preferred venue for this practice (Chakravarty, Gulen & Mayhew, 2004).

As hypothesised by Grossman (1986) and Hsu, Chiang and Liao (2013), there is a dearth of research examining information leakage in commodity markets. Much of the focus has examined the market reaction following macro-economic news announcements (Elder, Miao & Ramchander, 2012; Fleming & Remelona, 1999; Gorton & Rouwenhorst, 2006; Hess, Huang & Niessen, 2008; Roache & Rossi, 2010). Exceptions include Giot (2003) and Luo et al. (2016), who found significant informational content in IV of returns in the New York Board of Trade cacao, coffee and sugar futures markets and the Shanghai gold futures market.

[3] When the level of liquidity between two markets is sufficiently different, traditional measures of price discovery (e.g., the Hasbrouck [1995] IS) capture a significant amount of noise, which downplays the contribution of the relatively illiquid market in price discovery.

1.4 Price Discovery

The incorporation of information into prices is commonly referred to as price discovery. This important process in capital markets leads to two questions: Where is information first impounded into prices? Which venue do traders prefer? To examine this, the existing literature mainly relies on two methods: Hasbrouck (1995) and Gonzalo-Granger (1995). In addition, an extension of these methods developed by Yan-Zivot (2010) and Putniņš (2013) were utilised. These methods rely on a vector error correction model (VECM), which is applied to two price series to estimate the leader in price discovery. The greatest concern in this work is about the relationship between options and the underlying asset. Empirically, a number of studies show that:

1) Options lead stocks in terms of returns, volume or order flow (e.g., Chan, Chung & Fong, 2002; Easley et al., 1998; Hu, 2014).
2) Economically meaningful fractions of price discovery occur in options relative to stocks (e.g., Chakravarty et al., 2004; Patel et al., 2018).
3) Abnormal levels of options volume, order imbalance and implied volatility occur prior to takeover, earnings, splits and other corporate announcements (e.g., Cao, Chen & Griffin, 2005; Chae, 2005; Ghargori, Maberly & Nguyen, 2017; Jayaraman, Frye & Sabherwal, 2001).

Such empirical findings are consistent with how informed traders utilise a number of attractive features of options relative to equities. These include leverage, downside protection and the ability to overcome short sale constraints (e.g., Black, 1975; Lin & Lu, 2016).

1.5 Liquidity and Information Asymmetry

Based on the increase of technology and the speed of information transmission, discussions arose that suggested that markets might become more efficient. Prominent research by Brogaard, Hendershott and Riordan (2013) and O'Hara (2015) agree that this discussion is not settled, but prices became more efficient. They found no evidence of market instabilities. In this context, and by definition, information asymmetry due to insider trading affects trading activity. Riordan et al. (2013) highlights that the process of how information is transposed into prices is not fully discussed. The authors found higher adverse selection costs during times of information asymmetry and significantly varying liquidity levels during these times.

Understanding and measuring liquidity in times of information asymmetry is important for market makers and investors regarding their trading costs and risks and for regulatory bodies for monitoring market quality and insider trading. The literature examines liquidity in many contexts; for example, measuring trading costs (Amihud, 2002; Chung & Zhang, 2014; Corwin & Schultz, 2012), the relationship between liquidity and asset pricing (Anthonisz & Putniņš, 2016; Brennan et al., 2012; Lo & Wang, 2006) and capturing the presence of informed traders (Back, Crotty & Lee, 2018; Brennan, Hu & Subrahmanyam, 2018; Patel et al., 2018). In particular, using trades from Schedule 13D filings as a proxy for private information, Collin-Dufresne and Fos (2015) found that liquidity measures do not capture informed trading or behave consistently with microstructure theory. The authors show that such a result is possible because informed traders strategically trade when market liquidity is high and prefer to trade using limit orders.

Another branch of the liquidity literature examines the effect of earnings, dividends, takeovers, splits or repurchases and other firm announcements upon liquidity (in addition to share prices and trading activity). This branch of literature has significantly developed following the seminal research of Ball and Brown (1968); using an event study framework, the authors reported that earnings numbers are useful to investors, seemingly a given concept in today's financial markets.

1.6 Summary

Exploring the way in which information enters prices is an important topic, as it controls values that affect every market participant. The focus of this thesis is on derivative markets that are gaining popularity and inherit an increasing amount of information but are only now being recognised in the empirical research. After the introduction, this thesis considers price discovery through four distinct chapters.

Chapter 2 provided the first examination – to the best of our knowledge - of informed options trading around a sample of 352 Food and Drug Administration (FDA) announcements from 166 US-listed firms between 1996 and 2016. Using IV spreads and options trading volume as proxies for informed trading, two findings were highlighted: 1) informed traders are aware of the timing of the upcoming announcement at least five days in advance and 2) at least some informed traders have knowledge of the finer details that affect the price impact of the announcement. These findings have implications for regulators, investors and relevant firms.

In Chapter 3, the behaviour of US commodity futures and options IV-based measures are examined as proxies for information leakage around macro-economic and commodity-specific news announcements between 2007 and 2017. In the three days preceding news releases, abnormal changes in the levels of futures options IV spreads and skew are reported. In addition, a statistically significant relationship between announcement date returns and abnormal changes in pre-announcement IV spreads and skew are reported. Holistically, these findings indicate that at least some investors are informed about the price impact of upcoming news announcements in seven commodity markets.

Chapter 4 investigates whether commodity futures or options markets play a more important role in the price discovery process in the six most-actively traded markets: crude oil, natural gas, gold, silver, corn and soybeans. Using new information leadership techniques, new evidence reports that, in recent times, both markets make a meaningful contribution to price discovery. However, on average, options lead futures in reflecting new information for most of these commodities. It was found that increased speculation, rather than hedging activity, in commodity derivatives is a key determinant of price discovery in the options markets.

Chapter 5 uses an event study methodology to investigate whether the low-frequency (LF) measures of liquidity can proxy for HF measures around earnings announcements (i.e., periods of information asymmetry). Market microstructure data availability has significantly improved and it is now possible to estimate liquidity measures at the nanosecond level. However, this level of data are unavailable in all markets and time periods and there is a significant cost and computational burden of HF data. It was found that the closing-percent-quoted-spread is the best proxy for the percent–cost HF benchmarks. In contrast, using cross-sectional, portfolio and individual time-series correlations, the most consistent LF proxies are the high–low impact (HLI) and closing-percent-quoted-spread-impact. However, the performance of these proxies weakens in the pre- and post-announcement periods around the earnings announcement.

Finally, Chapter 6 wraps up this work and draws conclusions from each chapter. The chapter briefly outlines the research contributions for information leakage and price discovery and gives guidance on future research endeavours.

Chapter 2: Informed Options Trading Around US FDA Announcements

2.1 Research Question

The FDA is a key regulator, which aims to protect the US public from the use of a wide range of products, valued in excess of USD2.4 trillion. These products include food, drugs, biologics, medical devices or products, electronic products, cosmetics, veterinary products and tobacco products.[4] Given the extremely important role of the FDA for consumers, firms and investors, little is known about the material information contained in FDA announcements. There are only a few dated studies that found a marginal stock market reaction to FDA approval announcements of between zero and two percent between 1962 and 2001 (Bosch & Lee, 1994; Chaney, Devinney & Winer, 1991; Sharma & Lacey, 2004; Torabzadeh, Woodruff & Sen, 1998). The most recent study by Sarkar and de Jong (2006) decomposed FDA announcements into approvals and rejections and found abnormal announcement returns of 3.5 percent and –3.6 percent, respectively, with amplified reactions for smaller firms. Further, drug-related news is the third-most regular news event that has resulted in prosecuted insider trading filed by the SEC and the Department of Justice. This is evidence that suggests that FDA announcements contain material information and that there is information leakage prior to the announcement date (Patel et al., 2018).

This research examines whether informed trading takes place in options markets prior to FDA announcements. This question is examined due to the importance of the FDA in its role in society, the limited analysis of FDA announcements (particularly in recent times), evidence of illegal insider trading prior to such announcements and a stringent battery of literature supporting options as an important venue for informed trading. This research makes a unique contribution to the literature, as it is the first examination of informed options trading around FDA announcements. The unscheduled nature of FDA announcements aids this research design, as it is possible to more cleanly capture whether investors have prior knowledge of the forthcoming announcement (e.g., inside information passed on by company insiders). In contrast, for scheduled announcements,

[4] See https://www.fda.gov for details.

it is difficult to disentangle informed trading from changes in uncertainty, speculation or the superior ability of some individuals to predict the information content of such news.

This study examines a sample of 352 FDA new drug approvals (NDA) and drug-related announcements (FDA announcements) from 166 US-listed firms between 1 January 1996 and 31 December 2016. The IV spread developed by Cremers and Weinbaum (2010) is used, which has also been applied in other studies (e.g., Chan et al., 2015; Gharghori et al., 2017; Hao, 2016), as the primary proxy is informed trading. This measure captures price pressures from informed options trades. In addition, this study examines total options trading volume and call and put volume as additional proxies for informed trading.

Using multivariate panel regressions, the first key finding is that significant increases in average IV spreads and total, call and put options volume (expressed as a percentage of shares outstanding) are reported in the five days immediately preceding the FDA announcement. This finding indicates that at least some traders have knowledge of the timing of the forthcoming announcement. The second key finding is obtained from cross-sectional regressions, which show a statistically significant and positive relationship between pre-announcement abnormal IV spreads and abnormal FDA announcement returns. Given the positive news content of FDA announcements, a similar relationship is reported between abnormal announcement returns and abnormal pre-announcement call volume. The cross-sectional analysis indicates that some traders are informed about the finer details, which affects the price impact of the announcement (e.g., positive or negative news, type of FDA announcement).

This study will be interesting to regulators, investors and the relevant pharmaceutical and related firms. To the best of our knowledge, it is the first to examine and report informed trading in options markets around FDA announcements. This will be concerning to regulators and investors regarding the fairness of markets and the potential negative effects this could have upon the willingness of investors to trade, thereby affecting market liquidity. Pharmaceutical and other related firms should also be concerned by the leakage of information and may want to revise their corporate governance practices; in particular, the individuals trusted with such material information. These findings are also informative to regulators, as they increase knowledge of informed trading strategies, specifically where such investors choose to trade.

These findings contribute to the growing body of literature, which reports the abnormal behaviour of returns, volume, open interest and IV spreads in stock and options markets around corporate announcements. For example, prior to earnings releases (Amin & Lee, 1997; Jin, Livnat & Zhang, 2012; Johnson & So, 2012; Roll et al., 2010), takeovers (Acharya & Johnson, 2010; Augustin, Brenner & Subrahmanyam, 2015; Barraclough, Robinson, Smith & Whaley, 2013; Cao et al., 2005; Chan et al., 2015; Jayaraman et al., 2001; Podolski, Truong & Veeraraghavan, 2013), divestitures (Augustin, Brenner, Hu & Subrahmanyam, 2015; Patel & Michayluk, 2016a), repurchases (Hao, 2016), stock splits (Gharghori et al., 2017), stock issues (Henry & Koski, 2010) and analyst reports (Christophe et al., 2010; Hayunga & Lung, 2014; Lin & Lu, 2015). Holistically, the literature suggests that informed trading in options markets is a pervasive problem.

2.2 Data and Sample

Daily US options data were extracted from the *OptionMetrics IvyDB US* database and corresponding stock data from the *CRSP* and *Compustat* databases between 1 January 1996 and 31 December 2016. For the same sample period, 2,252 announcements from the FDA drug database were obtained (accessible via https://www.accessdata.fda.gov).

Table 2.1 summarises the sample selection filters. FDA announcement data were merged with *OptionMetrics*/*CRSP* data by ticker and company name, which reduced the sample to 898 announcements. Finally, announcements with missing abnormal IV spreads on any day in the five-day pre-announcement period [–5, –1] prior to the FDA announcement were excluded.[5] The final sample consists of 352 announcements from 166 companies during the 21-year sample period.

< Table 2.1 here >

The time series of FDA announcements are shown in Table 2.2. There is a sufficient number of FDA announcements per year (70+). After applying selection filters, there was an average of 17 announcements per year. The number of announcements in the sample fluctuated between a maximum of 30 in 2014 and a minimum of six in 1999. Most

[5] If this filter is omitted, the main results from implied volatility spreads are similar, using a sample of 898 FDA announcements.

announcements (257/352 or 73%) were given a 'standard' classification by the FDA; 95 announcements were given a 'priority' classification. Priority announcements are given for a 'drug that demonstrates the potential to provide a significant improvement in the safety or effectiveness of the treatment, diagnosis, or prevention of a serious or life-threatening condition from a drug that does not demonstrate such a potential' (FDA MAPP 6020.3 Rev. 2 Review Designation Policy).

< Table 2.2 here >

Table 2.3 shows a breakdown of the different classification types of FDA announcements. The FDA has 13 categories of announcements. For the final sample of 352 announcements, the most common type was 'new molecular entity' news (105/352 or 30%). A significant majority of announcements were standard announcements (266/352 or 76%) and were made by pharmaceutical companies with a Standard Industrial Classification (SIC) code 2834 (270/352 or 77%).[6]

< Table 2.3 here >

2.3 Informed Trading Measures

The primary informed trading measure is the implied volatility spread (IVS). Cremers and Weinbaum (2010) explained that widening IVS can proxy for informed trading. For example, when investors possess positive news, they may buy call options and sell put options. Given the positive relationship between option prices and implied volatility, this will increase the price spread and the IVS between call and put prices. A similar argument applies when investors have negative information. The open interest-weighted IVS for each stock day (it) was calculated:

$$IVS_{it} = \sum_{j=1}^{N_{it}} w_{jt}^{i} \left(IV_{jt}^{i,Call} - IV_{jt}^{i,Put}\right), \tag{2.1}$$

where $IV_{jt}^{i,Call}$ and $IV_{jt}^{i,Put}$ is the implied volatility for call and put options, N_{it} is the total number of put call pairs per stock day and w_{jt}^{i} is the average of the open interest between the call and put options for each put call pair. The implied volatility obtained from the

[6] Extended NDA classification code definitions can be accessed in the FDA manual of policies and procedures (MAPP) 5018.2.

OptionMetrics database is used, which is calculated using a binomial tree to account for discrete dividend payments and the early exercise premium. When constructing the IVS, the following options were excluded:

1) daily average quoted spreads greater than 50 percent
2) zero open interest
3) absolute delta smaller than 0.02 and greater than 0.98
4) incomplete data for put call pairs.

Similar results were obtained using a volume (rather than open interest) weighting to compute IVS.

In addition, options volume was used as a secondary proxy to capture informed trading. For each stock day, OSO was calculated, which is daily traded options volume scaled by shares outstanding. Call and put volume were also separately examined in which $\_Call$ (OSO_Put) is traded call (put) volume scaled by shares outstanding.

$$OSO_{it} = \frac{Option\ volume_{it}}{Shares\ outstanding_{it}}, \tag{2.2}$$

$$OSO\_Call_{it} = \frac{Call\ volume_{it}}{Shares\ outstanding_{it}}, \tag{2.3}$$

$$OSO\_Put_{it} = \frac{Put\ volume_{it}}{Shares\ outstanding_{it}}. \tag{2.4}$$

The abnormal values of each informed trading measure ($AMetric_{ie}$) were calculated by taking the difference between the daily trading measure during the event period [–10,+10] around the FDA announcement and the average value of the trading measure during the benchmark period [–60,–11]:

$$AMetric_{ie} = Metric_{ie} - \frac{1}{50}\sum_{t=-60}^{t=-11} Metric_{it}\,, e \in [-10, +10], \tag{2.5}$$

where $Metric_{ie}$ is either IVS, OSO, OSO_Call or OSO_Put. The FDA announcement date is day 0 and the pre-announcement period is [–5, –1].

2.4 Descriptive Statistics

Table 2.4 reports descriptive statistics. This analysis focuses on larger stocks by construction (i.e., stocks that have listed options). This is evident, as the average market capitalisation is over $18 billion. Consistent with prior studies, option-trading volume is considerably smaller than stock volume. For example, in the five days prior to the announcement, the mean stock volume (SVOL) (option volume, OSO) scaled by shares outstanding is 2.34 percent (0.05%). The average (IV) in the days prior to an FDA approval is approximately 41 percent, indicating a significant amount of future volatility. The average IVS is positive, indicating more options trading on positive, rather than negative, information. Consistent with such trading on positive news, the average abnormal return (ARET [0]) is approximately one percent.

< Table 2.4 here >

2.5 Univariate Plots of Abnormal Trading and Price Measures

Figure 2.1 plots the abnormal behaviour of IVS, OSO, IV and stock returns during the period [–10,+10] around the FDA announcement. Dotted lines indicate 90 percent confidence intervals. Panel A shows that the abnormal IVS is positive during the [–9,–2] period, reaching a maximum of approximately 0.5 percent on day –8, before slightly declining and rising again on day 0 to 0.7 percent. Following the announcement, abnormal IVS largely remains below 0 percent.

Panel B shows that abnormal OSO becomes positive on day –9 (0.02%) and continues to increase until day 0, reaching a maximum value of approximately 0.08 percent. Abnormal OSO on day –1 is more than twice as large as abnormal OSO during the period [–9, –2]. Post-announcement, abnormal OSO fluctuates by approximately 0.02 percent until day +9.

< Figure 2.1 here >

Abnormal levels of IV reported in Panel C show a similar pattern to abnormal OSO prior to the announcement date, increasing from approximately 1 percent on day –10 to 3 percent on day –1. When the FDA announcement is revealed to the market, abnormal IV quickly falls below 0 percent, consistent with the announcement reducing the level of uncertainty between investors.

Panel D shows the behaviour of abnormal returns. In general, abnormal returns fluctuate by approximately zero percent pre and post-announcement. However, on the announcement date, abnormal returns peak at one percent, which is consistent with the FDA announcement containing material information. Given that, by definition, this analysis focuses on larger stocks (which have liquid listed options), the abnormal returns reported from FDA announcements is smaller than similar prior studies, which do not impose such filters on their sample.[7] The lower abnormal announcement returns (and lower potential trading profits) may reduce the likelihood of pre-announcement informed trading. However, the subsequent analysis reports evidence supporting the presence of informed trading prior to FDA news.

Comprehensively, the abnormal behaviour of IVS, OSO, IV and stock returns provide initial evidence that at least some traders are informed about the upcoming FDA announcement. The following sections use multivariate analysis to more formally investigate informed trading.

2.6 Abnormal Implied Volatility Spreads

Following the univariate results, multivariate panel regressions are used to control for other factors and examine whether some traders are informed about the timing of the FDA announcement. It is anticipated that if some traders are informed about the timing of the announcement, the IVS should be abnormally elevated in the days preceding the FDA announcement. Using stock day observations during the window [–60,0], the following panel regressions are estimated:

$$IVS_{it} = \beta_0 + \beta_1 EventDV_{it} + \textstyle\sum_j \gamma_j \, Controls_{it} + \varepsilon_{it}, \quad (2.6)$$

[7] The abnormal FDA announcement returns are regressed on a dummy variable equal to one if the stock has listed options equal to zero. If the stock does not have listed options, it can be found that the abnormal announcement returns for stocks with listed options is approximately 0.5 percent smaller than for stocks without listed options (t-statistic = –2.22).

where $EventDV$ is a dummy variable equal to one during the pre-announcement period [–5, –1] and is equal to 0 during the benchmark period [–60, –11]. The following control variables are also included: $LagARET_{it}$ is the one–day lagged stock return minus the value-weighted *CRSP* index return, $|LagARET|_{it}$ is the absolute value of $LagARET_{it}$, $LagIVS_{it}$ is the one–day lagged value of IVS and $LagOSO_{it}$ is the one–day lagged options volume scaled by shares outstanding. In addition, event-fixed effects and cluster standard errors are included by month.

Table 2.5 Model 1 reports a positive and significant relationship between IVS and the $EventDV$ (t-statistic = 2.42). The $EventDV$ coefficient estimate indicates that the average IV spread increases by 0.226 percent in the five days prior to the FDA announcement from its benchmark level. Similar conclusions are reached in Models 2 and 3, in which an increasing number of control variables are included.[8] Using IVS as a proxy for informed trading suggests that at least a fraction of traders is informed about the timing of the upcoming FDA announcement.

< Table 2.5 here >

2.7 Informed Trading Prior to FDA Announcements

Following from these findings, this study investigates whether, in addition to knowledge of the timing of the announcement, traders are informed about the finer details that affect the price impact of the announcement (e.g., the priority status of the new drug). If some traders have such knowledge, then a positive relationship between abnormal IVS during the pre-announcement period and the announcement date returns is expected to be observed. This hypothesis is tested using the following cross-sectional regressions:

$$ARET_{it} = \beta_0 + \beta_1 AIVS[-5,-1]_{it} + \textstyle\sum_j \gamma_j \, Controls_{it} + \varepsilon_{i,t}, \qquad (2.7)$$

where $ARET_{i,t}$ is the abnormal return on the announcement date, which is calculated as the stock return minus the value-weighted CRSP index return and $AIVS_{it}$ is the average abnormal IV spread during the pre-announcement period [–5,–1]. Based on studies that found that, relative to stock prices, options play a minor role in the price discovery

[8] The adjusted R^2 ranges between 17 percent and 19 percent.

process, the following stock market control variables are included: $ARET[-5,-1]_{it}$ is the difference between stock returns and the value-weighted CRSP index return during the pre-announcement period [–5,–1], $SVOL[-5,-1]_{it}$ is the average SVOL scaled by shares outstanding during the pre-announcement period [–5,–1] and $Vola[-5,-1]_{it}$ is average stock return volatility during the pre-announcement period [–5,–1]. In addition, the following control variables are included: $PastReturn_{it}$ is the difference between stock returns and the value-weighted CRSP index calculated during the period [–40,–11], $StdReturn_{it}$ is the standard deviation of returns calculated over the period [–210,–11], $MktCap_{it}$ is the natural logarithm of market capitalisation of the firm as of the end of the fiscal year prior to the announcement and $PriorityDV_{it}$ is a dummy variable equal to one if the announcement is classified as priority and zero if the announcement is classified as standard. In addition, type (there are 13 different classification types), year fixed effect and cluster standard errors are included by firm.

Table 2.6 reports the cross-sectional regression findings. The key variable of interest is $AIVS[-5,-1]_{it}$. Consistent across Models 1 to 3, a positive (regression coefficient of 0.22) and significant (t-statistics of approximately 2) relationship between abnormal pre-announcement IVS and FDA announcement returns is reported. Using the $AIVS[-5,-1]_{it}$ regression coefficient in Model 3, a one standard deviation increase in $AIVS[-5,-1]_{it}$ (3.746) is associated with an estimated increase in $ARET_{i,t}$ by 0.809 percent (i.e., 0.216 x 3.746). An increase in $ARET_{i,t}$ by 0.809 percent is economically meaningful, as it captures 83 percent of the mean $ARET_{i,t}$ (i.e., 0.809%/0.973%), or 21 percent of its standard deviation (i.e., 0.809%/3.921). A significant proportion of the variation in $ARET_{it}$ is explained by the regression models with adjusted R^2 values between 11 percent and 14 percent. Using abnormal IVS as a proxy for informed trading, evidence is found that suggests that at least some traders are informed about the finer details of the FDA news, which affects its share price impact.

< Table 2.6 here >

2.8 Robustness

The findings reported in Sections 3.3. and 3.4. are robust to numerous specifications. First, the panel regression specifications in Eq. (2.6) are replicated. Instead, the $EventDV$

is defined as a dummy variable equal to one during the pre-announcement period [–10,–1] ([–3,–1]) and the findings are reported in Models 1 to 3 (4 to 6) of Table 2.7 Panel A. The motivation for these alternative definitions of the pre-announcement windows are based on the behaviour of IVS prior to FDA announcements, which are illustrated in Figure 2.1. These conclusions remain unchanged; average IV spreads significantly increase by 0.17 percent (0.238%) in the ten days (three days) prior to the FDA announcement from its benchmark level.

Second, the cross-sectional regression findings specified in Eq. (2.7) are robust if the average IVS during the periods [–10,–1] and [–3,–1,] is calculated. Table 2.7 Panel B reports such findings. The magnitude of the $AIVS[-10,-1]_{it}$ ($AIVS[-3,-1]_{it}$) regression coefficients in Model 3 (6.0) is approximately 1.6 (0.7) times larger than the respective coefficient in Table 2.6.

< Table 2.7 here >

Third, this research investigates how different characteristics, which affect the degree of informed trading, affect panel and cross-sectional findings. Using market capitalisation and whether the stock pays dividends, this research examines whether informed traders prefer smaller stocks when there is a larger information asymmetry between investors. Further, this study asks if informed traders prefer to trade strategically using liquid options, or options with a time to maturity that expires soon after the FDA announcement. Consistent with the characteristics of informed trading reported in prior studies (e.g., Easley et al., 1998), it was found that the results reported in Table 2.5 and Table 2.6 are stronger for stocks with listed options that are smaller by market capitalisation, do not pay dividends, have shorter time to maturity and are more liquid.[9]

2.9 Options Trading Volume

To examine whether these findings are specific to using IV spreads, options trading volume is used as a secondary proxy for informed trading. Similar to Eq. (2.6), panel regressions are run, except that IVS are replaced with options volume scaled by shares

[9] Robustness results are not reported for brevity.

outstanding (OSO) to examine if traders are informed about the FDA announcement timing:

$$OSO_{it} = \beta_0 + \beta_1 EventDV_{it} + \sum_j \gamma_j \, Controls_{it} + \varepsilon_{it}. \quad (2.8)$$

Table 2.8 Model 1 shows that the $EventDV_{it}$ variable is positive and significant, indicating that relative options volume is 3.92 percent larger in the five-day period prior to the FDA announcement relative to normal trading periods. If control variables are included, similar results are obtained in Models 2 and 3, in which options volume is 3.95 percent and 3.57 percent larger in the days preceding the announcement.

< Table 2.8 here >

OSO_{it} is decomposed into two components: call volume scaled by shares outstanding (OSO_Call_{it}) and put volume scaled by shares outstanding (OSO_Put_{it}). Using similar regression specifications detailed in Eq. (2.8), Table 2.9 Model 3 shows that call volume significantly increases by approximately 2.14 percent from the benchmark to the pre-announcement period. Similarly, Table 2.10 Model 3 shows that put volume significantly increases by approximately 1.43 percent from the benchmark to the pre-announcement period. Therefore, the significant increases in options volume observed prior to the FDA announcement are largely driven by increases in call volume (i.e., 2.14%/3.57% or 60%). Using options, call and put volume as measures of informed trading provide further evidence that some traders are aware of the date and time that FDA announcements are made.

< Table 2.9 here >

< Table 2.10 here >

Following these findings, similar cross-sectional regressions are used, as detailed in Eq. (2.7). However, the focus is on average abnormal OSO, rather than abnormal IVS, which investigates whether some individuals are privy to the contents of the FDA announcement:

$$ARET_{it} = \beta_0 + \beta_1 AOSO[-5,-1]_{it} + \sum_j \gamma_j \, Controls_{it} + \varepsilon_{i,t}. \quad (2.9)$$

Table 2.11 reports the results. In Models 2 and 3, which include a number of control variables, it was found that abnormal pre-announcement options volume is a significant determinant of the FDA announcement return. A one standard deviation increase in $AOSO[-5,-1]_{it}$ (0.022) is associated with an estimated increase in $ARET_{i,t}$ by 0.12 percent. This represents approximately 12 percent of the mean $ARET_{i,t}$. Relative to the IVS and $AIVS[-5,-1]$ results reported in Tables 2.5 and 2.6, the adjusted R^2 values of the regression models using OSO and $AOSO[-5,-1]$ are larger.

< Table 2.11 here >

From Figure 2.1 and the descriptive statistics, this research reports that FDA abnormal announcement returns are positive. As a result, it is anticipated that, as average FDA announcements contain positive news, investors will trade call, rather than put options, to maximise their returns. Therefore, call options volume will contain more information about announcement returns. Similar to the specifications in Table 2.11, this hypothesis is investigated by examining the relationship between $ARET_{it}$ and $AOSO\_Call[-5,-1]_{it}$ ($AOSO\_Put[-5,-1]_{it}$) in Table 2.12 (see Table 2.13).

< Table 2.12 here >
< Table 2.13 here >

Consistent with the hypothesis, it was found that abnormal levels of pre-announcement call volume (regression coefficient (t-statistic) of 9.516 (1.80) in Model 2 and 9.468 (2.11) in Model 3) significantly explain FDA abnormal announcement returns. Relative to changes in abnormal IV spreads, the economic magnitude of changes in abnormal options volume upon abnormal announcement returns is smaller. A one standard deviation increase in $AOSO\_Call[-5,-1]_{it}$ (0.014) is associated with an estimated increase in $ARET_{i,t}$ by 0.13 percent. This represents approximately 14 percent of the mean $ARET_{i,t}$. The findings report an insignificant relationship between pre-announcement abnormal put volume and announcement returns. These findings are consistent with some traders being informed about the details of the FDA announcements that affect stock price impact. Such traders utilise call options because, on average, FDA announcements contain positive news. As such, call options are informative about the abnormal announcement return.

2.10 Summary

To the best of our knowledge, this paper is the first to report evidence of informed trading in options markets prior to FDA announcements. The sample consists of 352 FDA announcements between 1996 and 2016. First, this shows that, using abnormal increases in IV spreads, options, call and put volumes, traders are informed about the timing of forthcoming FDA announcements. Second, it was found that such abnormal increases in IV spreads and options volume (in particular, call volume due to the positive nature of FDA announcements) is informative about the abnormal announcement return. Not only are some traders informed about the timing of the announcement, but they are also informed about the finer details, which dictate the price impact of the news. The unscheduled nature of FDA announcements suggests that these findings support the assertion that individuals have access to private information, rather than simply predicting or having a superior ability to process public information.

This study will be interesting to a number of parties, including regulators, investors and relevant firms. These findings add to the growing number of studies that document informed trading prior to corporate announcements in stock and options markets. Regulators should be concerned about the pervasive leakage of information. If some investors are privy to inside information (i.e., inequity in trading), then this can have negative effects on investors' willingness to trade and the liquidity of markets. Firms that make FDA announcements should be concerned about information leakage and may want to examine their corporate governance practices; in particular, the individuals that are trusted with such material information.

Figure 2.1: Abnormal trading behaviour around FDA announcements

This figure illustrates the average abnormal IV spreads (Panel A), abnormal options volume scaled by shares outstanding (Panel B), abnormal IV (Panel C) and abnormal stock returns (Panel D) around FDA announcements. Abnormal values of each trading or price measure were calculated as the difference between the daily trading or price measure during the event period [–10,+10] and the average of the trading or price measure during the benchmark period [–60,–11]. Dashed lines indicate the 90 percent confidence intervals. The horizontal axis expresses days relative to the announcement day 0. The sample comprises 352 announcements from 166 firms between 1 January 1996 and 31 December 2016.

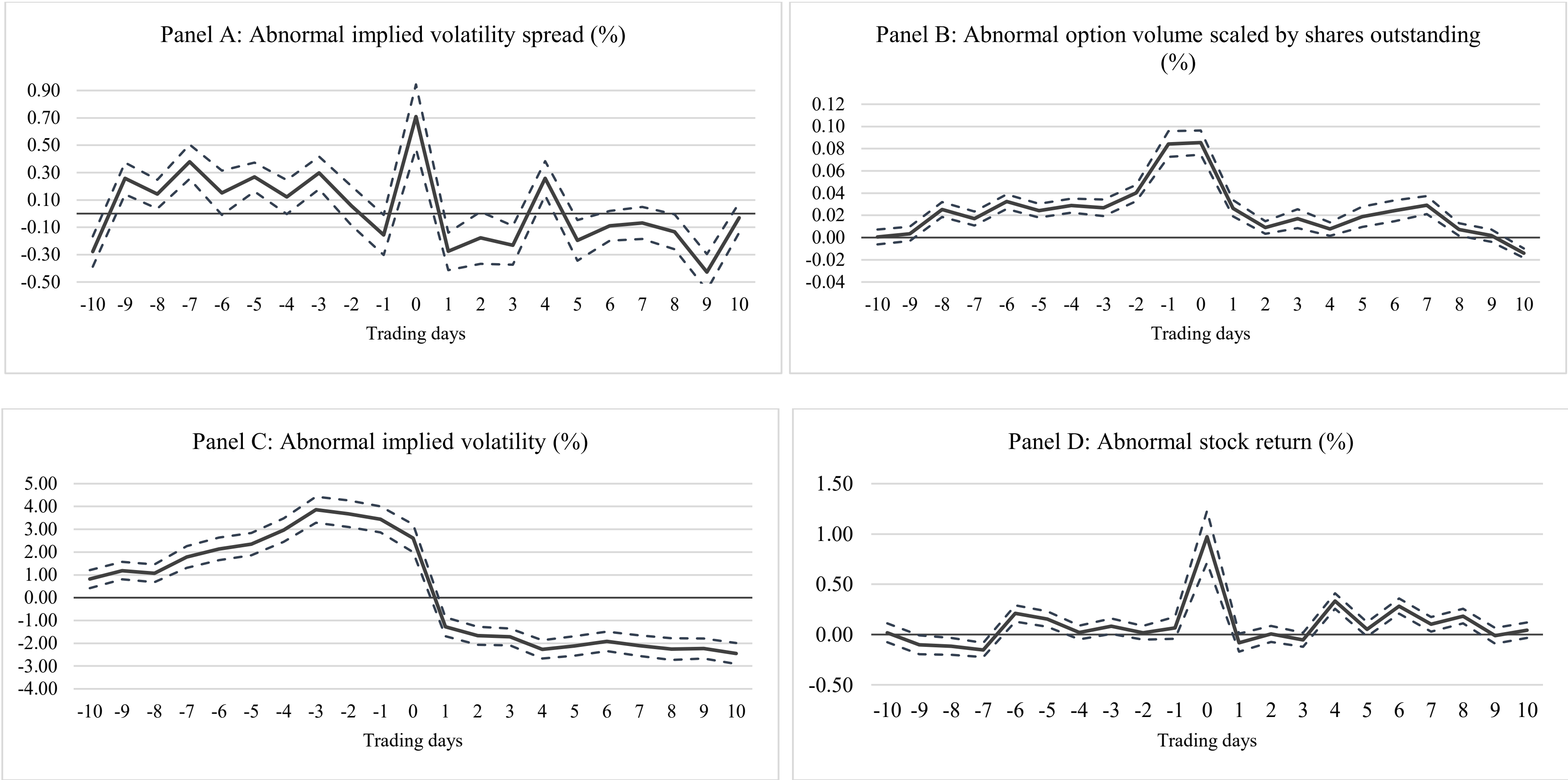


Table 2.1: Sample selection filters

This table reports the sample selection filters applied to the sample of FDA announcements between 1 January 1996 and 31 December 2016. FDA announcements are obtained from the FDA website. Daily options data are obtained from *OptionMetrics* and corresponding stock data from the *CRSP* and *Compustat* databases. Abnormal *IVS* is the average abnormal options IV spread during the pre-announcement period [–5,–1]. The FDA internally determines priority and standard classifications. Priority announcements are for drugs that demonstrate the potential to provide a significant improvement in the medical practice.

	Total	Priority	Standard
(1) Total number of FDA announcements between 1 January 1996 and 31 December 2016	2,252	437	1,815
(2) Keeping announcements with corresponding data in the *CRSP* database	1,385	262	1,123
(3) Keeping announcements with corresponding data in the *OptionMetrics* database	898	165	733
(4) Keeping announcements with non-missing abnormal *IVS* values during the pre-announcement period [–5,–1].	352	95	257

Table 2.2: FDA announcements over time

This table reports the number of FDA announcements between 1 January 1996 and 31 December 2016 for the total number of 2,252 FDA announcements and for 352 FDA announcements in the sample. The FDA internally determines standard and priority classifications. Priority announcements are for drugs that demonstrate the potential to provide a significant improvement in the medical practice.

Year	Total	Standard	Priority	Total	Standard	Priority
	2,252 announcements			352 announcements		
1996	143	118	25	17	10	7
1997	137	112	25	8	5	3
1998	92	78	14	13	9	4
1999	82	69	13	6	1	5
2000	106	91	15	10	8	2
2001	79	64	15	14	10	4
2002	78	64	14	16	15	1
2003	79	61	18	20	15	5
2004	102	90	12	22	11	11
2005	77	60	17	11	7	4
2006	111	87	24	29	21	8
2007	73	55	18	21	16	5
2008	78	59	19	17	15	2
2009	85	68	17	18	15	3
2010	84	66	18	21	19	2
2011	89	77	12	18	14	4
2012	89	81	8	14	11	3
2013	133	111	22	18	15	3
2014	106	87	19	30	19	11
2015	225	173	52	21	15	6
2016	204	163	41	8	6	2
Total	2,252	1,834	418	352	257	95

Table 2.3: Classification of FDA announcements

This table reports the number and classification of FDA announcements between 1 January 1996 and 31 December 2016 for the total number of 2,252 FDA announcements and for 352 FDA announcements in the sample. The FDA internally determines standard and priority classifications. Priority announcements are for drugs that demonstrate the potential to provide a significant improvement in the medical practice. In cases in which an FDA announcement is attributed to multiple classifications, the announcement is attributed to the first mentioned classification. *Rx to OTC switch* describes the change of a drug's status from requiring a practitioner prescription due to its toxicity or potential harm to the permission of OTC sales. Drugs classified as 'efficacy' are performance-enhancing supplements. Extended classification code definitions can be accessed in the FDA MAPP 5018.2. Pharmaceutical companies are classified by SIC code 2834.

Classification	Total	Standard	Priority	Total	Standard	Priority
	2,252 announcements			352 announcements		
1) New molecular entity	543	321	222	105	50	55
2) New active ingredient	43	36	7	9	4	14
3) New dosage form	768	673	95	5	5	0
4) New combination	220	187	33	0	0	0
5) New formulation or new manufacturer	288	260	28	1	1	0
6) New indication	111	95	16	100	76	15
7) Drug already marketed without approved NDA	44	38	6	3	3	0
8) Partial *Rx to OTC switch*	31	25	6	29	24	5
9) New indication submitted as distinct NDA, consolidated with original NDA after approval	44	37	7	38	34	4
10) New indication submitted as distinct NDA (not consolidated)	19	11	8	12	11	1
11) Efficacy	3	2	1	2	1	1
12) Medical gas	19	19	0	6	6	0
13) No classification	119	119	0	42	42	0
Total	2,252	1,823	429	352	257	95
Total (pharmaceutical companies only)	1,011	798	213	250	186	64

Table 2.4: Descriptive statistics of options and stock variables

This table shows descriptive statistics of options and stock-trading variables. All variable definitions are contained in Appendix A. The sample comprises 352 announcements from 166 firms between 1 January 1996 and 31 December 2016.

Variable	Mean	Median	Standard deviation
$ARET$ [0] (%)	0.973	0.187	3.921
IVS $[-5, -1]$ (%)	–0.268	–0.175	1.444
IVS $[-60, -11]$ (%)	–0.493	–0.300	3.760
$AIVS$ $[-5, -1]$ (%)	0.118	0.121	0.163
IV $[-5, -1]$ (%)	41.309	32.686	28.702
OSO (%)	0.117	0.038	0.268
OSO $[-5, -1]$ (%)	0.053	0.004	0.214
$AOSO$ $[-5, -1]$ (%)	0.041	0.029	0.022
$SVOL$ $[-5, -1]$ (%)	2.343	0.202	1.351
PRC ($)	32.467	22.110	39.707
$MktCap$ ($\$$ $millions$)	18,028.340	1,189.090	47,966.320
$StdReturn$ (%)	0.033	0.025	0.031

Table 2.5: Abnormal behaviour of IV spreads prior to FDA announcements

This table reports coefficient estimates from the following panel regressions of daily volatility spreads (IVS_{it}) on various independent variables using stock day observations during the window [–60,0]:

$$IVS_{it} = \beta_0 + \beta_1 EventDV_{it} + \sum_j \gamma_j \, Controls_{it} + \varepsilon_{it},$$

where $EventDV$ is a dummy variable equal to one during the pre-announcement period [–5,–1] and is equal to 0 during the benchmark period [–60,–11]. All variable definitions are contained in Appendix A. Event-fixed effects and cluster standard errors are included by month. The sample comprises 352 announcements from 166 firms between 1 January 1996 and 31 December 2016. *t*-statistics are reported in parentheses. ***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent level.

	Model 1	Model 2	Model 3
$Intercept$	–0.022	–0.021	–0.017
	(213.37)***	(–66.92)***	(–8.52)***
$EventDV$	0.226	0.224	0.216
	(2.42)**	(2.30)**	(2.40)**
$LagARET$		–0.046	0.008
		(–2.25)**	(0.34)
$\lvert LagARET \rvert$			–0.014
			(–0.48)
$LagIVS$			0.165
			(2.84)***
$LagOSO$			–0.003
			(–0.35)
$Event\ FE$	Yes	Yes	Yes
N	18,824	18,824	18,824
$Adj. R^2$	0.17	0.17	0.19

Table 2.6: Abnormal IV spreads and announcement date returns

This table reports coefficient estimates from the following cross-sectional regressions of abnormal announcement day zero returns ($ARET_{it}$) on various independent variables using stock day observations:

$$ARET_{it} = \beta_1 AIVS[-5,-1]_{it} + \sum_j \gamma_j \, Controls_{it} + \varepsilon_{i,t},$$

where $AIVS[-5,-1]_{it}$ is the abnormal IV spread averaged during the pre-announcement period [–5,–1]. All variable definitions are contained in Appendix A. FDA classification type fixed effects (13 different classification types), year fixed effects and cluster standard errors are included by firm. The sample comprises 352 announcements from 166 firms between 1 January 1996 and 31 December 2016. t-statistics are reported in parentheses.***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent level.

	Model 1	Model 2	Model 3
$Intercept$	0.021	0.023	0.060
	(2.59)***	(2.73)***	(1.52)
$AIVS[-5,-1]$	0.222	0.214	0.216
	(2.47)**	(2.18)**	(2.06)**
$ARET[-5,-1]$		–0.118	–0.142
		(–0.44)	(–0.53)
$SVOL[-5,-1]$		–0.045	–0.316
		(–0.21)	(–1.10)
$Vola[-5,-1]$		–0.208	–0.276
		(–1.34)	(–1.68)*
$PastReturn$			–0.111
			(–0.06)
$StdReturn$			0.126
			(0.49)
$MktCap$			–2.098
			(–1.08)
$PriorityDV$	0.001	0.001	0.001
	(0.26)	(0.23)	(0.22)
$Type\ FE$	Yes	Yes	Yes
$Year\ FE$	Yes	Yes	Yes
N	352	352	352
$Adj. R^2$	0.11	0.12	0.14

Table 2.7: Abnormal IV spreads and announcement date returns (robustness)

This table replicates the analysis in Tables 2.5 and 2.6 using different definitions of the periods of interest. Panel A reports coefficient estimates from the following panel regressions of daily volatility spreads (IVS_{it}) on various independent variables using stock day observations during the window [–60,0]:

$$IVS_{it} = \beta_0 + \beta_1 EventDV_{it} + \sum_j \gamma_j \, Controls_{it} + \varepsilon_{it},$$

where $EventDV$ is a dummy variable equal to one during the pre-announcement period [–10,–1] ([–3,1]) in Models 1 to 3 (Models 4 to 6) and is equal to 0 during the benchmark period [–60,–11]. All variable definitions are contained in Appendix A. Event-fixed effects and cluster standard errors are included by month.

Panel B reports coefficient estimates from the following cross-sectional regressions of abnormal announcement day 0 returns ($ARET_{it}$) on various independent variables using stock day observations:

$$ARET_{it} = \beta_1 AIVS[-X,-1]_{it} + \sum_j \gamma_j \, Controls_{it} + \varepsilon_{i,t},$$

where $AIVS[-X,-1]_{it}$ is the abnormal IV spread averaged during the pre-announcement period [–10,–1] ([–3,–1]) in Models 1 to 3 (Models 4 to 6). FDA classification type fixed effects (13 different classification types), year fixed effects and cluster standard errors are included by firm. The sample comprises 352 announcements from 166 firms between 1 January 1996 and 31 December 2016. *t*-statistics are reported in parentheses.***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent level.

Panel A	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
$Intercept$	–0.020	–0.019	–0.016	–0.022	–0.021	–0.018
	(–121.19)***	(–50.42)***	(–9.57)***	(–283.63)***	(–72.68)***	(–8.25)***
$EventDV$	0.191	0.185	0.170	0.229	0.238	0.238
	(2.17)**	(1.99)**	(2.04)**	(2.20)**	(2.20)**	(2.32)**
$LagARET$		–0.048	0.007		–0.044	0.010
		(–2.17)**	(0.33)		(–2.11)**	(0.45)
$\|LagARET\|$			–0.007			–0.016
			(–0.29)			(–0.53)
$LagIVS$			0.179			0.161
			(3.05)***			(2.74)***
$LagOSO$			–0.002			–0.002
			(–0.36)			(–0.18)
$Event\ FE$	Yes	Yes	Yes	Yes	Yes	Yes
N	20,184	20,184	20,184	18,128	18,128	18,128
$Adj. R^2$	0.18	0.18	0.21	0.17	0.17	0.19
$Intercept$	0.020	0.020	0.045	0.020	0.022	0.042
	(2.45)**	(2.23)**	(1.12)	(2.45)**	(2.63)***	(0.97)
$AIVS[-X,-1]$	0.315	0.347	0.352	0.137	0.130	0.146
	(2.88)***	(3.01)***	(2.70)***	(2.19)**	(2.01)**	(2.18)**
$ARET[-X,-1]$		0.314	0.354		–0.194	–0.185
		(0.93)	(0.97)		(–1.31)	(–1.22)
$SVOL[-X,-1]$		0.047	–0.114		–0.127	–0.303
		(0.19)	(–0.41)		(–0.57)	(–0.94)
$Vola[-X,-1]$		–0.172	–0.203		0.024	–0.009
		(–0.88)	(–0.98)		(0.18)	(–0.06)
$PastReturn$			0.671			–1.144
			(0.35)			(–0.57)
$StdReturn$			0.054			0.137
			(0.21)			(0.53)
$MktCap$			–1.380			–1.288
			(–0.68)			(–0.59)
$PriorityDV$	0.001	0.001	0.001	0.002	0.001	0.002
	(0.28)	(0.12)	(0.10)	(0.29)	(0.23)	(0.29)
$Type\ FE$	Yes	Yes	Yes	Yes	Yes	Yes
$Year\ FE$	Yes	Yes	Yes	Yes	Yes	Yes
N	352	352	352	352	352	352
$Adj. R^2$	0.12	0.13	0.14	0.10	0.12	0.13

Table 2.8: Abnormal behaviour of options volume prior to FDA announcements

This table reports coefficient estimates from the following panel regressions of daily option volumes scaled by shares outstanding (OSO_{it}) on various independent variables using stock day observations during the window [–60,0]:

$$OSO_{it} = \beta_0 + \beta_1 EventDV_{it} + \sum_j \gamma_j \, Controls_{it} + \varepsilon_{it},$$

where $EventDV$ is a dummy variable equal to one during the pre-announcement period [–5,–1] and equal to 0 during the benchmark period [–60,–11]. All variable definitions are contained in Appendix A. Event-fixed effects and cluster standard errors are included by month. The sample comprises 352 announcements from 166 firms between 1 January 1996 and 31 December 2016. *t*-statistics are reported in parentheses. ***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent level.

	Model 1	Model 2	Model 3
$Intercept$	0.104 (131.29)***	0.100 (72.47)***	0.062 (13.65)***
$EventDV$	3.918 (5.50)***	3.945 (5.48)***	3.568 (5.42)***
$LagARET$		0.236 (3.06)***	0.017 (0.20)
$\|LagARET\|$			0.530 (6.03)***
$LagIVS$			–0.076 (–2.20)**
$LagOSO$			0.350 (6.56)***
$Event\ FE$	Yes	Yes	Yes
N	18,473	18,473	18,473
$Adj. R^2$	0.53	0.53	0.55

Table 2.9: Abnormal behaviour of call volume prior to FDA announcements

This table reports coefficient estimates from the following panel regressions of daily call volume scaled by shares outstanding (OSO_Call_{it}) on various independent variables using stock day observations during the window [–60,0]:

$$OSO\_Call_{it} = \beta_0 + \beta_1 EventDV_{it} + \sum_j \gamma_j\, Controls_{it} + \varepsilon_{it},$$

where $EventDV$ is a dummy variable equal to one during the pre-announcement period [–5,–1] and equal to zero during the benchmark period [–60,–11]. All variable definitions are contained in Appendix A. Event-fixed effects and cluster standard errors are included by month. The sample comprises 352 announcements from 166 firms between 1 January 1996 and 31 December 2016. t-statistics are reported in parentheses. ***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent level.

	Model 1	Model 2	Model 3
$Intercept$	0.056	0.053	0.032
	(110.37)***	(63.25)***	(8.8)***
$EventDV$	2.286	2.299	2.135
	(5.03)***	(4.98)***	(4.72)***
$LagARET$		0.154	0.039
		(2.75)***	(0.54)
$\|LagARET\|$			0.352
			(5.16)***
$LagIVS$			0.009
			(0.27)
$LagOSO\_Call$			0.152
			(4.00)***
$Event\ FE$	Yes	Yes	Yes
N	18,473	18,473	18,473
$Adj.R^2$	0.45	0.45	0.46

Table 2.10: Abnormal behaviour of put volume prior to FDA announcements

This table reports coefficient estimates from the following panel regressions of daily put volume scaled by shares outstanding (OSO_Put_{it}) on various independent variables using stock day observations during the window [–60,0]:

$$OSO\_Put_{it} = \beta_0 + \beta_1 EventDV_{it} + \sum_j \gamma_j \, Controls_{it} + \varepsilon_{it},$$

where $EventDV$ is a dummy variable equal to one during the pre-announcement period [–5,–1] and equal to zero during the benchmark period [–60,–11]. All variable definitions are contained in Appendix A. Event-fixed effects and cluster standard errors are included by month. The sample comprises 352 announcements from 166 firms between 1 January 1996 and 31 December 2016. t-statistics are reported in parentheses. ***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent level.

	Model 1	Model 2	Model 3
$Intercept$	0.048	0.048	0.030
	(121.58)***	(67.96)***	(13.30)***
$EventDV$	1.632	1.646	1.434
	(4.58)***	(4.62)***	(4.57)***
$LagARET$		0.082	–0.021
		(1.68)*	(–0.55)
$\|LagARET\|$			0.179
			(3.29)***
$LagIVS$			–0.085
			(–2.94)***
$LagOSO\_Put$			0.198
			(8.93)***
$Event\ FE$	Yes	Yes	Yes
N	18,473	18,473	18,473
$Adj. R^2$	0.41	0.41	0.43

Table 2.11: Abnormal options volume and announcement date returns

This table reports coefficient estimates from the following cross-sectional regressions of abnormal announcement day 0 returns ($ARET_{it}$) on various independent variables using stock day observations:

$$ARET_{it} = \beta_1 AOSO[-5,-1]_{it} + \sum_j \gamma_j \, Controls_{it} + \varepsilon_{i,t},$$

where $AOSO[-5,-1]_{it}$ is the abnormal options volume scaled by shares outstanding averaged during the pre-announcement period [–5,–1]. All variable definitions are contained in Appendix A. FDA classification type fixed effects (13 different classification types), year fixed effects and cluster standard errors are included by firm. The sample comprises 352 announcements from 166 firms between 1 January 1996 and 31 December 2016. *t*-statistics are reported in parentheses.***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent level.

	Model 1	Model 2	Model 3
$Intercept$	–0.009	–0.006	–0.138
	(–0.46)	(–0.32)	(–2.23)**
$AOSO[-5,-1]$	3.055	5.218	5.632
	(1.10)	(1.69)*	(1.97)**
$ARET[-5,-1]$		0.142	0.118
		(0.48)	(0.40)
$SVOL[-5,-1]$		–0.310	–0.178
		(–0.93)	(–0.48)
$Vola[-5,-1]$		–0.283	–0.317
		(–0.77)	(–0.84)
$PastReturn$			–6.950
			(–2.01)**
$StdReturn$			1.001
			(2.68)***
$MktCap$			6.290
			(2.22)**
$PriorityDV$	–0.008	–0.008	–0.008
	(–1.22)	(–1.18)	(–1.22)
$Type\ FE$	Yes	Yes	Yes
$Year\ FE$	Yes	Yes	Yes
N	341	341	341
$Adj. R^2$	0.14	0.16	0.21

Table 2.12: Abnormal call volume and announcement date returns

This table reports coefficient estimates from the following cross-sectional regressions of abnormal announcement day 0 returns ($ARET_{it}$) on various independent variables using stock day observations:

$$ARET_{it} = \beta_1 AOSO\_Call[-5,-1]_{it} + \sum_j \gamma_j \, Controls_{it} + \varepsilon_{i,t},$$

where $AOSO\_Call[-5,-1]_{it}$ is the abnormal call volume scaled by shares outstanding averaged during the pre-announcement period [–5,–1]. All variable definitions are contained in Appendix A. FDA classification type fixed effects (13 different classification types), year fixed effects and cluster standard errors are included by firm. The sample comprises 352 announcements from 166 firms between 1 January 1996 and 31 December 2016. *t*-statistics are reported in parentheses.***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent level.

	Model 1	Model 2	Model 3
$Intercept$	–0.005	0.001	–0.128
	(–0.26)	(–0.02)	(–2.08)**
$AOSO\_Call[-5,-1]$	6.314	9.516	9.468
	(1.34)	(1.80)*	(2.11)**
$ARET[-5,-1]$		0.141	0.119
		(0.48)	(0.39)
$SVOL[-5,-1]$		–0.348	–0.203
		(–1.09)	(–0.53)
$Vola[-5,-1]$		–0.284	–0.312
		(–0.84)	(–0.91)
$PastReturn$			–6.741
			(–1.90)*
$StdReturn$			0.972
			(2.82)***
$MktCap$			6.097
			(2.17)**
$PriorityDV$	–0.009	–0.008	–0.008
	(–1.29)	(–1.22)	(–1.24)
$Type\ FE$	Yes	Yes	Yes
$Year\ FE$	Yes	Yes	Yes
N	341	341	341
$Adj. R^2$	0.15	0.17	0.22

Table 2.13: Abnormal put volume and announcement date returns

This table reports coefficient estimates from the following cross-sectional regressions of abnormal announcement day 0 returns ($ARET_{it}$) on various independent variables using stock day observations:

$$ARET_{it} = \beta_1 AOSO\_Put[-5,-1]_{it} + \sum_j \gamma_j \, Controls_{it} + \varepsilon_{i,t},$$

where $AOSO\_Put[-5,-1]_{it}$ is the abnormal put volume scaled by shares outstanding averaged during the pre-announcement period [–5,–1]. All variable definitions are contained in Appendix A. FDA classification type fixed effects (13 different classification types), year fixed effects and cluster standard errors are included by firm. The sample comprises 352 announcements from 166 firms between 1 January 1996 and 31 December 2016. *t*-statistics are reported in parentheses.***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent level.

	Model 1	Model 2	Model 3
$Intercept$	–0.010	–0.009	–0.141
	(–0.51)	(–0.48)	(–2.24)**
$AOSO\_Put[-5,-1]$	0.747	2.381	3.695
	(0.17)	(0.45)	(0.74)
$ARET[-5,-1]$		0.148	0.124
		(0.50)	(0.42)
$SVOL[-5,-1]$		–0.129	0.004
		(–0.37)	(0.01)
$Vola[-5,-1]$		–0.224	–0.260
		(–0.56)	(–0.63)
$PastReturn$			–6.910
			(–1.94)*
$StdReturn$			0.990
			(2.48)**
$MktCap$			6.264
			(2.16)**
$PriorityDV$	–0.007	–0.007	–0.007
	(–1.08)	(–1.06)	(–1.11)
$Type\ FE$	Yes	Yes	Yes
$Year\ FE$	Yes	Yes	Yes
N	341	341	341
$Adj. R^2$	0.14	0.14	0.20

Appendix A – Chapter 2

Definition of variables

Variable	Definition
$AIVS[X,Y]$	Abnormal IV spread averaged over days [X,Y]. Abnormal IV spread is defined in Eq. (2.5).
$AOSO[X,Y]$	Abnormal options volume scaled by shares outstanding averaged over days [X,Y]. Abnormal options volume scaled by shares outstanding is defined in Eq. (2.5).
$AOSO\_Call[X,Y]$	Abnormal call options volume scaled by shares outstanding averaged over days [X,Y]. Abnormal call options volume scaled by shares outstanding is defined in Eq. (2.5).
$AOSO\_Put[X,Y]$	Abnormal put options volume scaled by shares outstanding averaged over days [X,Y]. Abnormal put options volume scaled by shares outstanding is defined in Eq. (2.5).
$ARET[X]$	Announcing a firm's stock day return minus the stock day return on the value-weighted *CRSP* market index on day [X].
$ARET[X,Y]$	Announcing a firm's stock return during days [X,Y] minus the return on the value-weighted CRSP market index during days [X,Y].
$EventDV$	A dummy variable equal to one if the day is in the event window and zero if the day is in the pre-event benchmark period.
$IV[X,Y]$	IV averaged over days [X,Y].
$IVS[X,Y]$	IVS averaged over days [X,Y]. IVS is defined in Eq. (2.1).
$LagIVS$	One–day lagged IVS. IV spread is defined in Eq. (2.1).
$LagOSO$	One–day lagged options volume scaled by shares outstanding. Options volume includes total options volume across all contracts multiplied by 100.

Variable	Definition
$LagARET$	One–day lagged announcing a firm's stock day return minus the stock day return on the value-weighted *CRSP* market index on day [X].
$MktCap$	The natural logarithm of market capitalisation as of the end of the fiscal year prior to the announcement.
$OSO[X,Y]$	Options volume scaled by shares outstanding averaged over days [X,Y]. Options volume includes total options volume across all contracts multiplied by 100.
$OSO\_Call[X,Y]$	Call options volume scaled by shares outstanding averaged over days [X,Y]. Call volume includes total call volume across all contracts multiplied by 100.
$OSO\_Put[X,Y]$	Put options volume scaled by shares outstanding averaged over days [X,Y]. Put volume includes total put volume across all contracts multiplied by 100.
$PastReturn$	The announcing firm's stock day return minus the stock day return on the value-weighted *CRSP* index calculated during the period [–40,–11] prior to the announcement date.
PRC	Closing share price averaged over days [X,Y].
$PriorityDV$	Dummy variable equal to one if the announcement is classified as priority and zero if the announcement is classified as standard by the FDA.
$RET[X]$	The announcing firm's stock day return on day [X].
$StdReturn$	Standard deviation of returns calculated over 200 trading days during the period [–210,–11] prior to the announcement.
$SVOL[X,Y]$	SVOL scaled by shares outstanding averaged over days [X,Y].
$Vola[X,Y]$	Stock return volatility averaged over days [X,Y].

Chapter 3: Information Leakage in Commodity Derivatives Around News Announcements

3.1 Research Question

Given the substantial body of literature supporting options as a venue for information leakage, the lack of information leakage research in commodity options markets and the significant growth in futures options trading, this research investigates and provides—to the best of our knowledge—the first comprehensive study of whether information leakage takes place in agriculture, energy and metals commodity futures options markets prior to news announcements.

This study examines the behaviour of IV-based measures of information leakage prior to macro-economic and commodity-specific news releases for the agriculture (corn, soybeans and wheat), energy (crude oil and natural gas) and metals (gold and silver) futures options markets between 2007 and 2017. This research focuses on these markets because they have the highest trading volume on the CME. Specifically, IV-based measures are used, as they have significant return predictability in the equity options markets. These are IV spread (Cremers & Weinbaum, 2010) and IV skew (Xing, Zhang & Zhao, 2010). Cremers and Weinbaum (2010) explained that the IV spread captures information leakage because traders buy (sell) calls and sell (buy) puts when they possess positive (negative) information. The effect on prices is directly reflected in IV and are thereby captured by the IV spread. In addition, following Xing, Zhang and Zhao (2010), the IV skew is used. IV skew captures the difference in IV between out-of-the-money (OTM) and at-the-money (ATM) options. An increase in the IV differential reflects a preference by informed traders for options with increased leverage.[10]

The methodology used in this research comprises three steps: a univariate analysis and two multivariate panel regression analyses. Univariate plots of the behaviour of IV-based measures indicate increased trading activity in the three-day period prior to macro-economic news and commodity-specific reports. In particular, this study reports increases in IV spread by 0.4 percent to 1 percent in the three days preceding positive news in crude

[10] Cheng and Xiong (2014) reached a similar conclusion.

oil, gold and silver markets. In addition, for all commodities, increases in IV skew are reported that reflect investors' preferences for enhanced leverage of OTM put options prior to negative news. After controlling for a number of market factors, the panel regression analysis supports the univariate findings. Statistically significant changes in IV spread are reported for five out of seven commodities in the days preceding the announcement as investors trade in anticipation of the upcoming news. Finally, a another set of panel regressions examining the relation between announcement date returns and abnormal pre-announcement changes in IV spread suggests that at least some traders are informed and familiar with details of news announcement (e.g., demand forecasts for natural gas), which determine the price impact of the announcement in all commodity markets.

These findings will be interesting to regulators, investors and firms that utilise commodities. This study reports evidence of information leakage in a number of commodity markets. This will be concerning to regulators and investors regarding the fairness of markets and potential negative effects this could have on investor's willingness to trade and market liquidity. These findings are also informative to regulators, as they increase knowledge of information leakage strategies; in particular, where and when such investors choose to trade.

3.2 Data and Methodology

This research examines agriculture, energy and metal futures options contracts with the highest trading volume traded on the CME as at 2017. These include corn (contract name: ZC), soybeans (contract name: ZS), wheat (contract name: ZW), crude oil (contract name: CL), natural gas (contract name: NG), gold (contract name: GC) and silver (contract name: SI). Options and underlying futures prices, volume, IV and open interest between 1 January 2007 and 31 March 2017 were obtained from *IVolatility*.[11] *IVolatility* data are utilised by more than 70,000 clients including the top 50 (30) investment banks (option market makers), hedge funds, the CBOE, the New York Stock Exchange and the Options

[11] *IVolatility* calculate the implied volatility for options on individual stocks using a binomial tree to account for discrete dividend payments and the early exercise premium.

Clearing Corporation and were also utilised in academic research (e.g., Mixon, 2009; 2011).

Following Augustin et al. (2015) and Hao (2016), the following options were excluded. Those with:

1) zero open interest
2) time to expiration of fewer than five days and more than 90 days
3) deep out of the money (DOTM) and deep in the money options (DITM)
4) option put-call pairs in which either the call or put option is missing.

Option moneyness is determined using the ratio of the underlying price to the strike price (Augustin, Brenner & Subrahmanyam, 2015; Xing et al., 2010):

$$Moneyness_t = \frac{Underlying\ Price_t}{Strike\ Price_t}, \tag{3.1}$$

where OTM options are defined by moneyness values between 0.8 and 0.95, ATM options have values between 0.95 and 1.05 and in the money options have values between 1.05 and 1.2. Option moneyness levels exceeding the extremes are classified as either DOTM or DITM.

All macro-economic and commodity-specific report announcement dates are obtained via *Bloomberg*. Table 3.1 lists and summarises the news events included in this study. Three macro-economic news events are studied that have had the largest economic effect on metal prices, including US consumer price index (CPI), US industrial production (IP) and US gross domestic product (GDP) announcements (Roache & Rossi, 2010). CPI announcements are released monthly by the US Bureau of Labor Statistics. IP is a measure of industrial output from the manufacturing, mining and utility sectors and is released on a monthly basis by the Federal Reserve Board of the US. The US Bureau of Economic Analysis releases quarterly GDP estimates.

< Table 3.1 here >

In addition, information leakage around commodity-specific news events issued by several government authorities is examined. Announcements are chosen that have the highest relevance to each commodity according to *Bloomberg*. *Bloomberg* determines relevance by the number of alert subscriptions for each announcement. For agriculture markets, the monthly 'World Agricultural Supply and Demand Estimates Report (WASDE)' issued by the US Department of Agriculture is used. The report includes estimates of supply and demand for major crops and livestock. For energy markets, the monthly 'Short-Term Crude' and 'Natural Gas Outlook' reports issued by the US Energy Information Administration are used. These reports include statistics and forecasts on energy prices, production, demand and exploration.[12]

For macro-economic and commodity-specific announcements, positive and negative news announcements are examined separately to avoid contrasting price effects and changes in implied volatility-based measures. A positive (negative) news announcement is defined when the difference between the actual announcement value and the median analyst forecast value (obtained from *Bloomberg*), scaled by the standard deviation of the forecasting error across analysts, is greater (less) than zero (Bernile, Hu & Tang, 2016; Elder et al., 2012).[13]

3.3 Implied Volatility-based Metrics

This research estimates two IV-based measures that have been found to predict future stock returns as proxies for information leakage in commodity futures options markets. Following Cremers and Weinbaum (2010), IV spread is estimated in Eq. (3.2):

$$IV\ Spread_t = w\left(IV_t^{Call} - IV_t^{Put}\right), \tag{3.2}$$

where $IV_{jt}^{i,Call}$ and $IV_{jt}^{i,Put}$ is the IV for call and put options computed by *OptionMetrics* and w_{jt}^{i} is the average of the open interest between the call and put option for each put call pair per trading day.

Using Eq. (3.3), IV skew (put) is estimated following Xing et al. (2010), which captures the increased leverage of OTM put options. In addition, an adjusted version, IV skew

[12] Announcements, which occur after the market is closed or on a weekend, are shifted to the next trading day.
[13] For each announcement, the standard deviation of the forecasting error across analysts is the standard deviation of the difference between the actual announcement value and the median analyst forecast value.

(call), is estimated, which captures informed traders' preference for the leverage of OTM call options:

$$IV\ Skew\ (put)_t = \left(IV_t^{OTMPut} - IV_t^{ATMCall}\right), \tag{3.3}$$

$$IV\ Skew\ (call)_t = \left(IV_t^{OTMCall} - IV_t^{ATMCall}\right), \tag{3.4}$$

where IV_t^{OTMPut}, $IV_t^{OTMCall}$ and $IV_t^{ATMCall}$ is the IV for OTM put, OTM call and ATM call options computed using *OptionMetrics*.[14] For example, when computing IV skew (put) for each trading day, for options with the same time to maturity, the average IV across OTM put options and across ATM call options (where moneyness is defined in Eq. [3.1]) is calculated. These averages are used to compute IV skew (put).

3.4 Event Study Methodology

In a similar vein to Hao (2016) and Kacperczyk and Pagnotta (2018), for agriculture, energy and metal commodities, the abnormal behaviour of each implied volatility measure around news announcements is examined using an event study methodology. For each announcement date (day [0]), the focus is on the behaviour of implied volatility measures during the three-day pre-announcement period [–3,–1]. The definition of the pre-announcement period is motivated by the scheduled earnings literature, which tends to find that the probability of information leakage is elevated in the one to three-day period prior to the announcement (Back et al., 2018; Brennan et al., 2018). A benchmark (or control) period is defined, which includes *k* days in the calendar quarter of the announcement, excluding all days corresponding to the event period [–3,0] of any announcements occurring in that calendar quarter. Days in each calendar quarter are used to account for calendar quarter seasonality in options and futures markets. In the case of metals, in which three different macro-economic news announcements are examined, in cases in which the event period [–3,0] of two or more announcements intersect, more recent announcement(s) are omitted from the analysis. The abnormal value of implied volatility proxies are calculated for information leakage by taking the difference between the daily implied volatility measure during event days [–3,0] and the average of the daily

[14] The implied volatility of ATM call options is used as the benchmark level of implied volatility because they are the most actively traded and liquid options contract.

implied volatility measures during the benchmark period (which consists of k days in each calendar quarter).

3.5 Summary Statistics and Univariate Analysis

Table 3.2 reports descriptive statistics regarding the trading activity of commodity futures, call and put option contracts during the sample period. The most actively traded commodities in terms of futures and options volume are corn, crude oil, natural gas and gold. For example, the average daily options volume for corn contracts is 35,524, which is relative to the daily futures volume of 130,996. The average daily option volume as a percentage of the average daily futures volume ranges from a minimum of 8 percent in silver markets to 27 percent in corn and soybeans, indicating that a reasonable amount of trading activity occurs in futures options. Comprehensively, futures returns are close to 0 percent, except for wheat, which has an average return of 0.68 percent. Gold has the lowest average IV of 21 percent, while future volatility for natural gas exceeds 48 percent. Generally, the remaining commodities have implied volatilities ranging between 20 percent and 35 percent.

< Table 3.2 here >

Table 3.3 reports descriptive statistics for IV spread and IV skew for each commodity calculated between 2007 and 2017. Corn, gold and silver show the largest mean IV spread values of approximately 0.5 percent, while both metals and crude oil have the largest IV skews, with values between 3 percent and 5.2 percent. Wheat and crude oil have negative average IV skew and IV spread values, respectively.

< Table 3.3 here >

As a starting point, each commodity is examined in terms of the abnormal behaviour of futures returns, IV spread and IV skew during the period [–3,0] around positive and negative news announcement dates.[15] Figure 3.1 illustrates this.

< Figure 3.1 here >

[15] Only plots during the period [–3,–1] are presented to ensure that sufficient non-event days are presented in the control period.

For each commodity, Panel A reports abnormal futures returns. In most cases, spikes in returns are observed on the announcement date, as news becomes public information. The largest announcement date returns of 0.6 percent and –0.8 percent occur in the soybeans and wheat markets, which suggests that such news announcements generally contain material information. Pre-announcement spikes in abnormal returns were noted prior to the announcement date. This is likely due to information leakage prior to the announcement dates.

The abnormal behaviour of IV spread and IV skew varies for each commodity prior to the news date. For example, the most prominent increases in abnormal IV spread prior to positive news events occur in crude oil (0.5% between days [–3,–1]), gold (0.4% on day –1) and silver (more than 1% between days [–3,–2]). Consistent with expectations, IV skew (put) is largely positive in the [–3,–1] period prior to negative news for all commodities. It is likely that this is caused by increases in the trading of OTM put options. Agriculture commodities reveal negative levels of abnormal IV skew (call) in the three days prior to positive news, compared to metals, which show positive levels.

It was also found that abnormal futures options volume is higher during the pre-announcement period for all commodities and for both positive and negative news events. Notably, abnormal volume reaches a maximum of 4,000 on day –1 for crude oil and exceeds 3,000 contracts for natural gas during the period [–3,–1]. However, pre-announcement abnormal volume is much closer to 0 for soybeans, wheat and silver commodities.

It should be noted that there are several mechanisms with which information is imputed into prices. The most obvious is that informed investors trade options and their trading pressure results in the revision of prices by market makers. If option volume is not driving price changes, then information may be revealed in prices and quotes via informed market makers (Hao, 2016). Another possibility is that investors trade the underlying futures contract, rather than futures options. The focus of this research is on options markets and it is not possible to examine such mechanisms using daily option data. As such, this avenue could be explored in future research, with more detailed data regarding the identity and trading of option market makers and futures data answering these questions.

3.6 Benchmark Panel Regressions

Using univariate analysis, Section 3.5 documents abnormal changes in IV (i.e., spreads and skew) around macro-economic and commodity-specific news announcements. In this section, using multivariate regression analysis with event-fixed effects, the question is examined of whether IV spread and IV skew during the pre-announcement period are significantly different to benchmark period values due to the impending news announcement or whether they can be explained by a number of market factors:

$$METRIC_{i,t} = \beta_0 + \beta_1 EventDV + \beta_2 Controls + \varepsilon, \tag{3.5}$$

where *METRIC* is IV spread, IV skew (put) or IV skew (call) during all days in the benchmark period and the pre-announcement period. *EventDV* is a dummy variable that is equal to 0 during the benchmark period (which consists of days in which there is no pre-announcement or announcement days during a calendar quarter) and is equal to 1 during the pre-event period [–3,–1]. The control variables include one–day lagged return (*Lag(Return)),* one–day lagged options volume *(Lag(OptionsVolume))* and one–day lagged options open interest *(Lag(OptionsOI))*. Due to the scheduled nature of the announcements examined in this study, pre-announcement abnormal changes in the IV-based measures that are examine can be caused by informed traders taking options positions, uninformed individuals making speculative bets or hedging activity.

The variable of interest in the regression framework is *EventDV*. This indicates whether IV metrics in the three days prior to the announcement date are significantly different relative to the benchmark period. For regressions focusing on IV spread, if information leakage is occurring (or speculative bets in the right direction of the news) for positive (negative) news, a positive (negative) coefficient on the *EventDV* variable consistent with IV spread being higher (lower) prior to the announcement is expected, as traders buy (sell) calls and sell (buy) put options. For regressions focusing on IV skew (put), if informed or speculative trading in the right direction is occurring, for positive (negative) news, a negative (positive) coefficient is expected on the *EventDV* variable that is consistent with IV skew (put) being lower (higher) prior to the announcement. This is because traders buy (sell) ATM calls and sell (buy) OTM put options. For regressions focusing on IV skew (call), for either positive or negative news, either a positive or a negative coefficient on the *EventDV* is expected, depending on the relative amounts of ATM and OTM call option trading. For example, for positive news, informed traders may utilised the increased leverage of OTM relative ATM call options, increasing IV skew (call), resulting in a positive coefficient on the *EventDV* variable. However, the opposite is also possible if informed traders prefer the liquidity of ATM options (Augustin et al., 2016).

Table 3.4 documents the findings. For each commodity, Panels A, B and C report regressions that examine IV spread, IV skew (put) and IV skew (call), respectively. Panel A reports abnormal changes in IV spread for five out of seven of the commodities that are examined. For positive (negative) news, significant increases (decreases) in IV spread (as captured by positive [negative] signs on the *EventDV*) are reported during the three-day period prior to the announcement for wheat and crude oil (natural gas, gold and silver). These findings are consistent with informed traders or speculators that bet in the same price direction as the upcoming news announcement. In addition, positive news is also reported, as abnormal decreases in IV spread resulting from hedging or speculative trading in the opposite direction of the news for natural gas, gold and silver. Such findings do not preclude at least some traders from being informed about positive news in the natural gas, gold and silver markets.

< Table 3.4 here >

Panel B reports abnormal increases in IV skew (put) prior to negative news in corn, soybean, natural gas, gold and silver markets. These abnormal increases in IV skew (put) are consistent with informed traders or speculators leveraging their returns by purchasing OTM put options in the presence of bad news. Abnormal increases in IV skew (put) prior to positive news announcements are reported in the wheat, crude oil, natural gas, gold and silver markets. These patterns are likely to be due to bets from uninformed traders. However, another possibility is the creation of protective put strategies (potentially by informed investors) through the purchase of OTM puts, in addition to the purchase of underlying futures contracts.[16]

Panel C reports abnormal changes in IV skew (call) across five commodities. For positive news regarding forecasted energy prices, production and demand, abnormal increases in IV skew (call) are reported for crude oil and natural gas, which is consistent with traders

[16] Protective put strategies are commonly used for investors to limit downside risk while retaining unlimited profit potential.

(who are potentially informed) preferring OTM call options relative to their ATM counterparts. In contrast, positive macro-economic news announcements result in traders purchasing more liquid ATM call options, indicating that options liquidity—or the ability to strategically trade—is important for their trading strategies. Similarly, for negative news relating to soybeans and silver, it is reported that traders also prefer trading ATM options.

3.7 Return-based Panel Regressions

Section 3.6 reported abnormal changes in the level of IV-based measures prior to news announcements for all seven commodity markets. Evidence was reported that was consistent with information leakage and speculative bets in both the same and opposite direction to the price direction of the news announcement. In this section, another regression analysis is conducted to examine whether such abnormal changes in these measures observed during the pre-event window are driven by informed traders who have knowledge of the details of the news announcements that determine its price impact (e.g., increases in demand for crude oil). The following regression is estimated:

$$RET_{i,t} = \beta_0 + \beta_1 AbnMETRIC[-3,-1] + \beta_2 Controls + \varepsilon, \tag{3.6}$$

where *RET* is the futures announcement date return. *AbnMETRIC* is the average abnormal IV spread, IV skew (put) or IV skew (call) during the period [–3,–1]. Control variables include abnormal return in the pre-event period (*Abn. Return (%) [–3,–1]*), average daily options volume (*options volume*) and average daily options open interest (*options open interest*). Year dummy variables are also included.

The variable of interest in the regression framework is *AbnMETRIC*. For regressions in which IV spread is the dependent variable, a positive coefficient is expected on the *AbnMETRIC* variable that is consistent with abnormal changes in the IV metrics (or information leakage activity) having a positive association with the announcement return. The same positive relation between *AbnMETRIC* and *RET* is expected for both positive and negative news. For positive (negative) news, informed traders will buy (sell) calls or sell (buy) put options. This increases (decreases) IV spread. Thus, a positive relation with positive (negative) announcement returns is anticipated. For regressions in which IV skew (put) is the dependent variable, a negative coefficient on the *AbnMETRIC* variable is expected to be consistent with abnormal changes in IV skew (put) having a negative association with the announcement return. The same negative relation is expected between *AbnMETRIC* and *RET* for both positive and negative surprises. For negative news, informed traders will buy OTM put options to take advantage of the increased leverage inherent in OTM relative to ATM options. This will cause the IV of OTM put options to increase, increasing IV skew (put). Thus, a negative relation with negative announcement returns is expected. In contrast, for positive news, informed traders will likely buy call options, causing the IV of ATM call options to increase, decreasing IV skew (put). Thus, a negative relation with positive announcement returns is expected. For regressions in which IV skew (call) is the dependent variable, a positive or negative coefficient could be expected on the *AbnMETRIC* variable depending on the relative amount of trading of OTM and ATM call options. For example, for positive news, informed traders may utilise the increased leverage of OTM relative to ATM call options, increasing IV skew (call). As such, a positive relation with positive announcement returns is anticipated. However, for positive news, it is possible that informed traders may purchase ATM relative to OTM options for liquidity or strategic reasons or depending on the price effect of the news (Augustin et al., 2016).

Table 3.5 presents the regression results. For each commodity, Panel A examines the relationship between announcement returns and abnormal IV spread. For each commodity, it is shown that pre-announcement abnormal IV spread is informative about the news announcement returns. For positive news, the *AbnMETRIC* variable has a positive and significant coefficient for announcements relating to corn, wheat, crude oil, natural gas and gold. For example, Models 1 to 3 report *AbnMETRIC* coefficient estimates of between 0.743 and 0.918 for natural gas and associated *t*-statistics of approximately 4.50. Similarly, for negative news, abnormal IV spread is positively associated with announcement returns for soybeans and silver. These findings are consistent with informed traders buying (selling) call options or selling (buying) put options prior to good (bad) news, in which either trade contributes to an increase (decrease) in IV spread.

< Table 3.5 here >

For each commodity, Panel B reports results in which *AbnMETRIC* is represented by IV skew (put). In general, the coefficient estimate of abnormal IV skew (put) is an insignificant determinant of the announcement date return, except in the energy market. In both crude oil and natural gas commodities, it was found that abnormal increases in IV skew (put) are associated with negative news about forecasted energy prices, production, demand or exploration, which is consistent with informed traders preferring trading OTM put options to maximise their returns. This causes the IV of OTM options to increase.

Finally, Panel C examines whether abnormal levels of IV skew (call) explain announcement date returns. The pre-announcement abnormal differential in IV between OTM and ATM call options is positively related to the returns associated with positive news in natural gas and gold markets, highlighting investors' increased demand for OTM options, which can maximise returns. In contrast, the differential in OTM and ATM implied volatilities could be explained by informed traders preferring to sell OTM call options prior to negative news in corn and soybeans and sell ATM call options prior to negative news in natural gas markets.

Section 3.6 reported abnormal changes in IV-based measures prior to news announcements. Such abnormal changes in IV spread and skew can occur as a result of information leakage, other investors making speculative bets prior to the scheduled news announcement or hedging activity. Thoroughly, the return-based panel regressions indicate that the behaviour of abnormal IV-based measures is informative about the magnitude and direction of the news announcement. Although it is not possible to formally distinguish between information leakage and speculation, in all cases in the regression models in which the *AbnMETRIC* variable is significant, the sign of the coefficient estimate is in the expected direction, which is consistent with information leakage. This analysis suggests that there is information leakage prior to macro-economic and commodity-specific news and that at least some information leakage takes place in commodity options markets.

3.8 Summary

This study provides one of the most comprehensive examinations of information leakage in agriculture, energy and metal commodity markets prior to macro-economic and commodity-specific news announcements. In particular, the most actively traded commodity options and futures contracts are examined for the underlying assets of corn, soybeans, wheat, crude oil, natural gas, gold and silver. The main analysis uses IV-based measures that have been used in a number of prior studies as proxies for information leakage, such as Augustin, Brenner & Subrahmanyam (2015), Cremers & Weinbaum (2010) and Xing, Zhang & Zhao (2010).

First, the initial evidence is presented via univariate plots of abnormally elevated levels of trading prior to commodity news announcements. Second, the abnormal behaviour of IV spread and IV skew, prior to news, is confirmed using panel regressions. This is present after controlling for several market factors. Third, using return-based panel regressions, a statistically significant relationship between abnormal announcement returns and abnormal changes in pre-announcement IV-based measures is reported in all seven commodity markets. These results indicate that at least some traders have finer knowledge regarding the price impact of the announcement. Given the significant advances in trading and technology over time, future research can examine the prevalence of information leakage in other futures options markets (e.g., interest rates).

Figure 3.1: Abnormal metrics around surprise announcements

This figure shows average abnormal IV spread, IV skew (put) and IV skew (call), abnormal implied volatilities, abnormal returns and abnormal options volumes (all variable definitions are provided in Appendix A) for corn, soybeans, wheat, crude oil, natural gas, gold and silver in the [–3,0] period around positive and negative surprise announcements. All variables are presented as percentages. The sample includes futures and options data between 2007 and 2017.

Corn

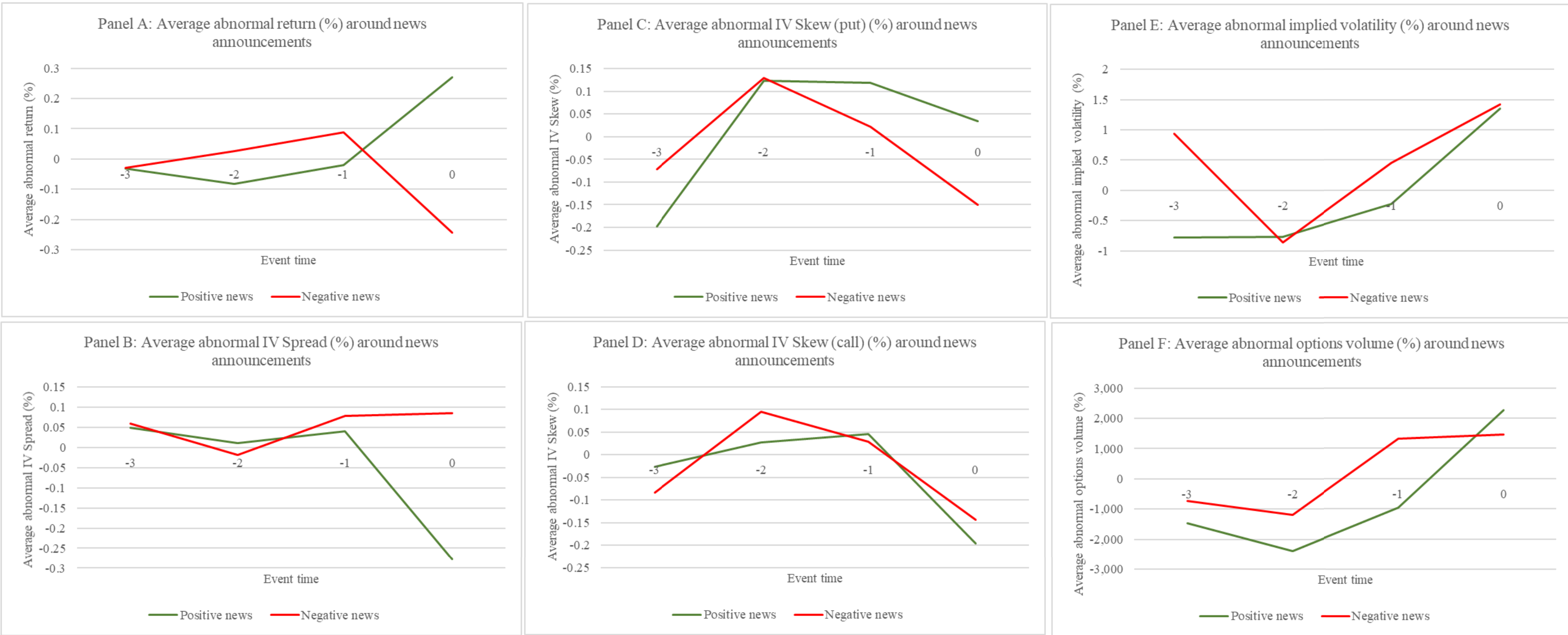


Figure 3.1: Abnormal metrics around surprise announcements (continued)

Soybeans

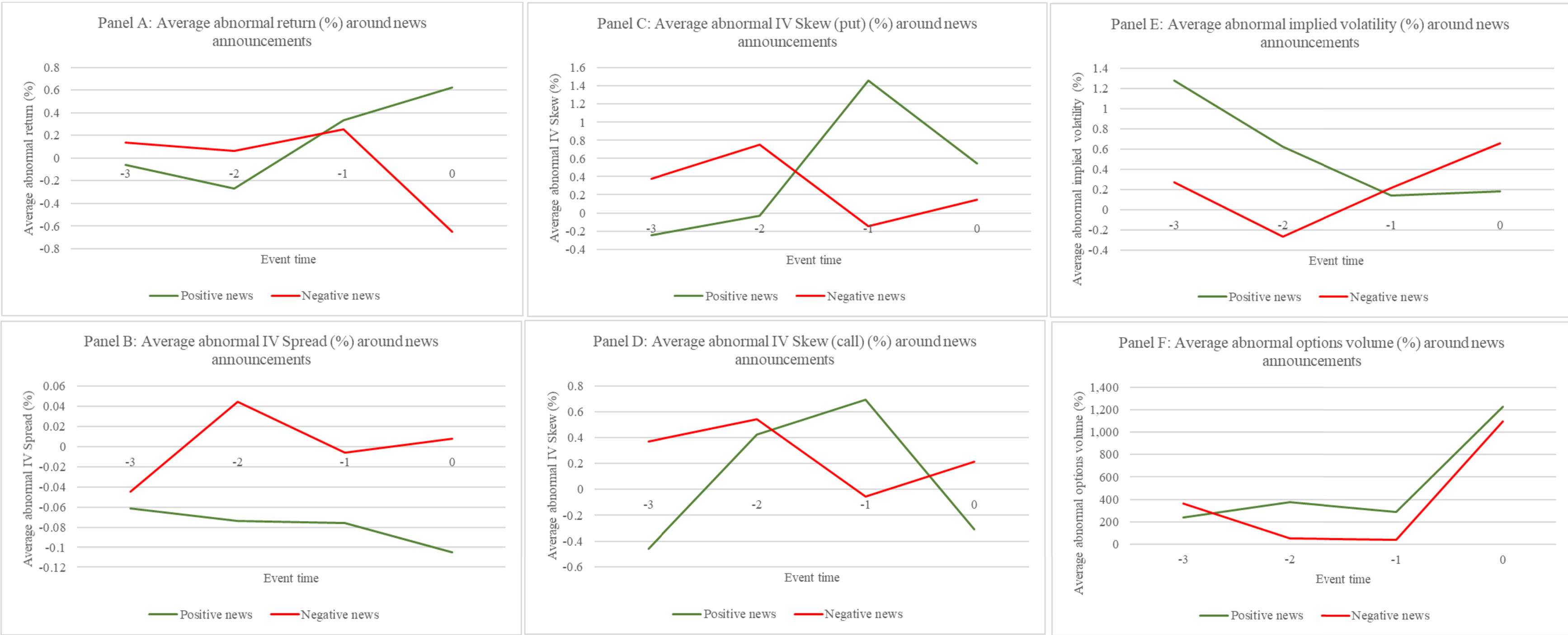


Figure 3.1: Abnormal metrics around surprise announcements (continued)

Wheat

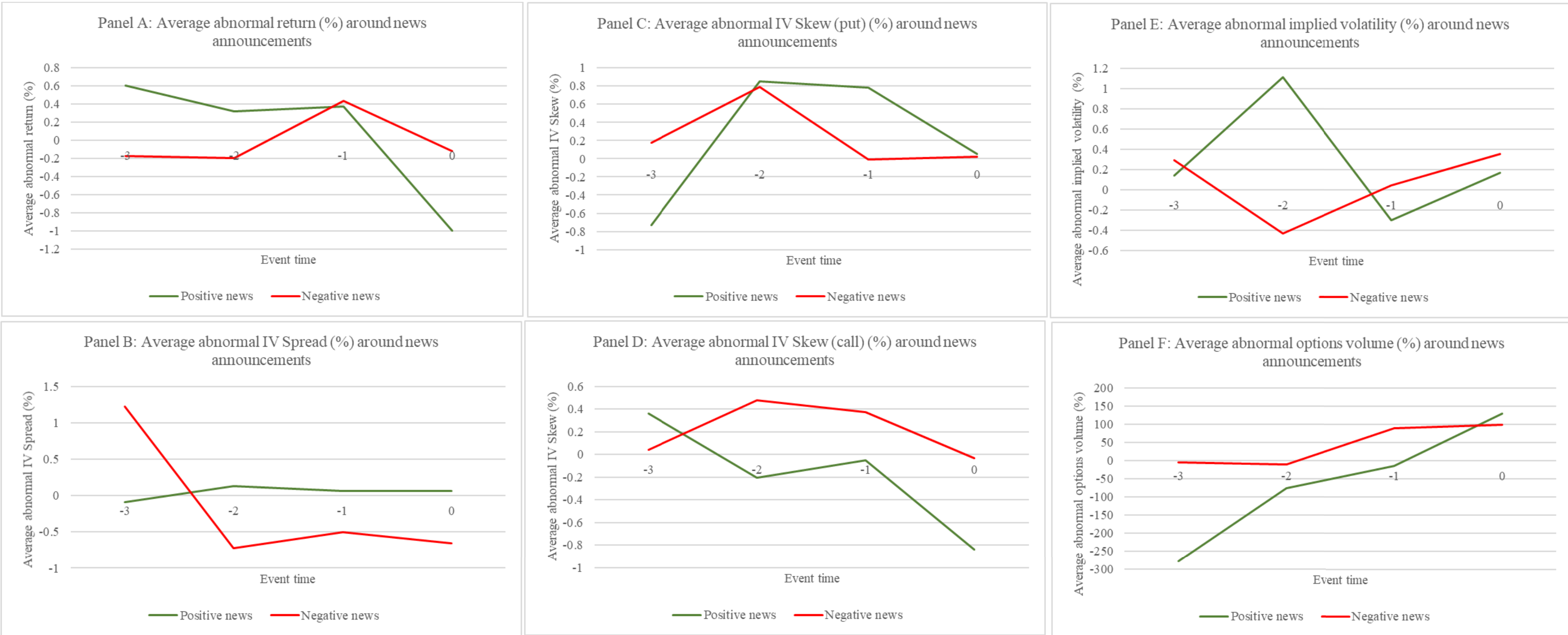


Figure 3.1: Abnormal metrics around surprise announcements (continued)

Crude oil

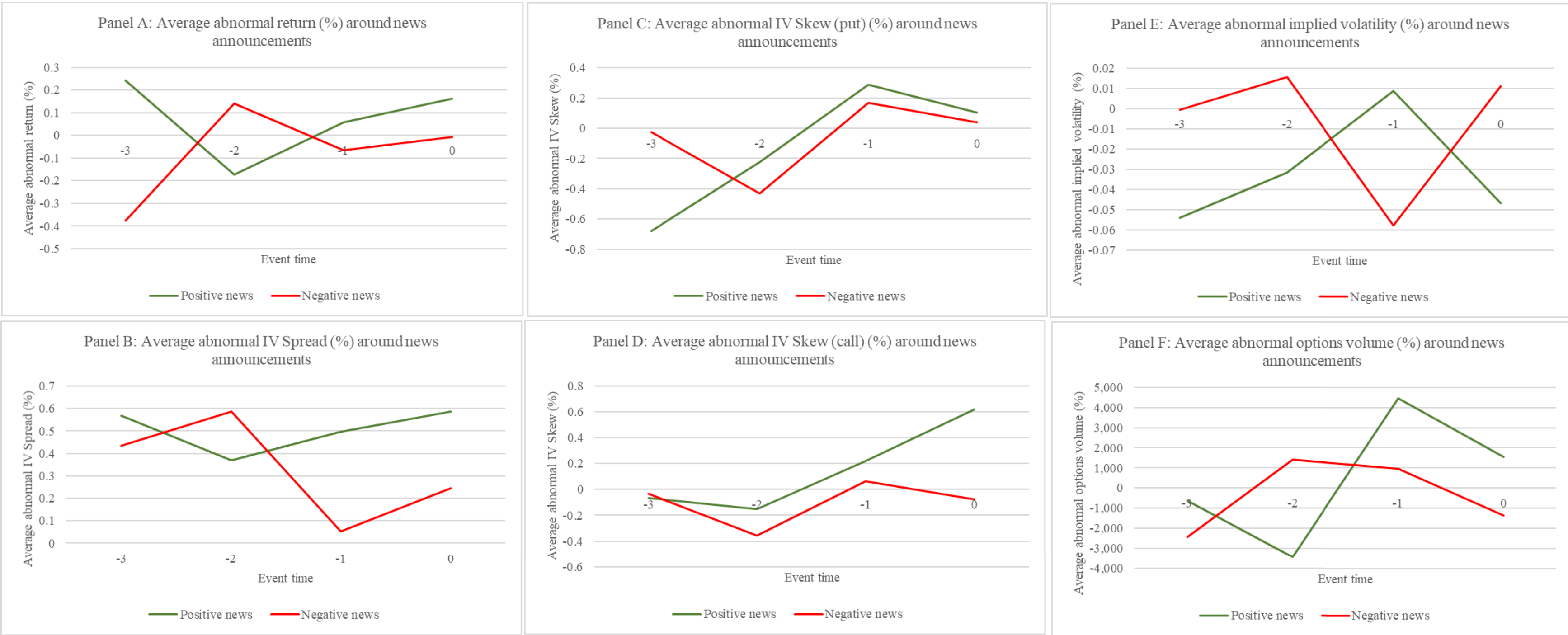


Figure 3.1: Abnormal metrics around surprise announcements (continued)

Natural gas

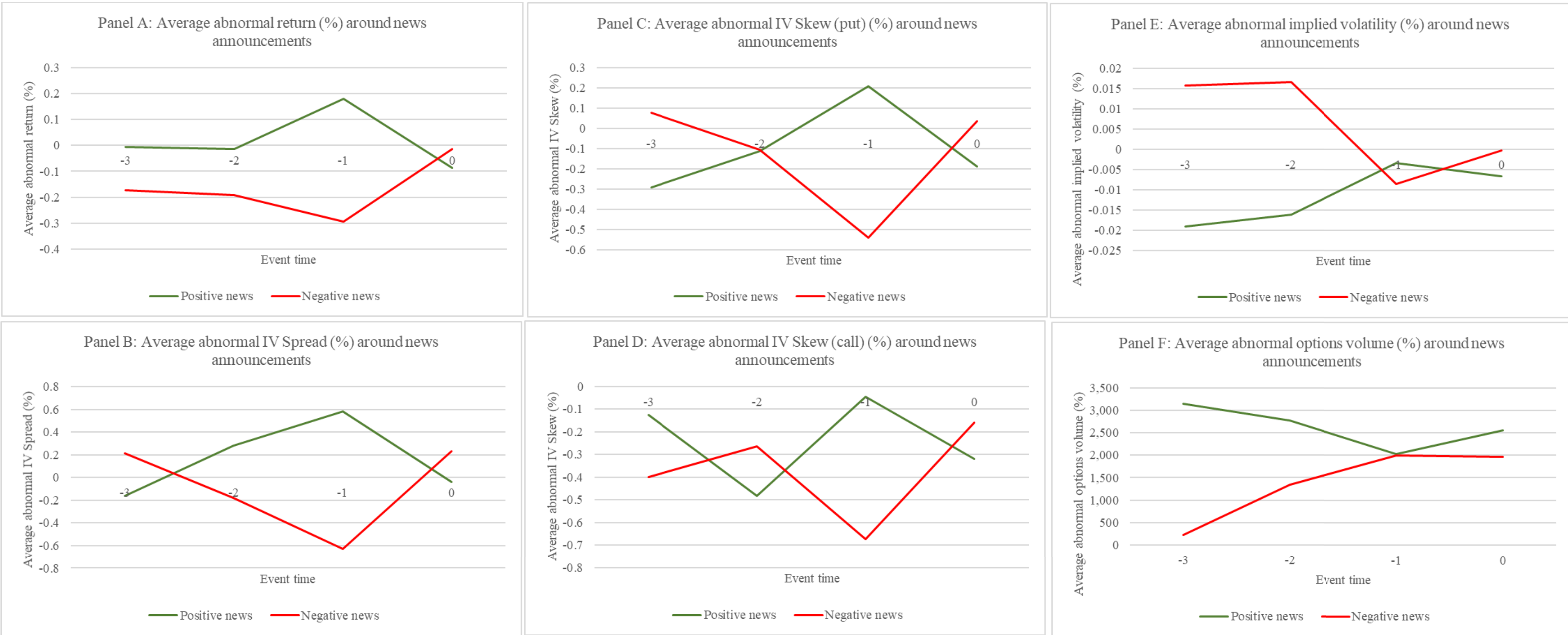


Figure 3.1: Abnormal metrics around surprise announcements (continued)

Gold

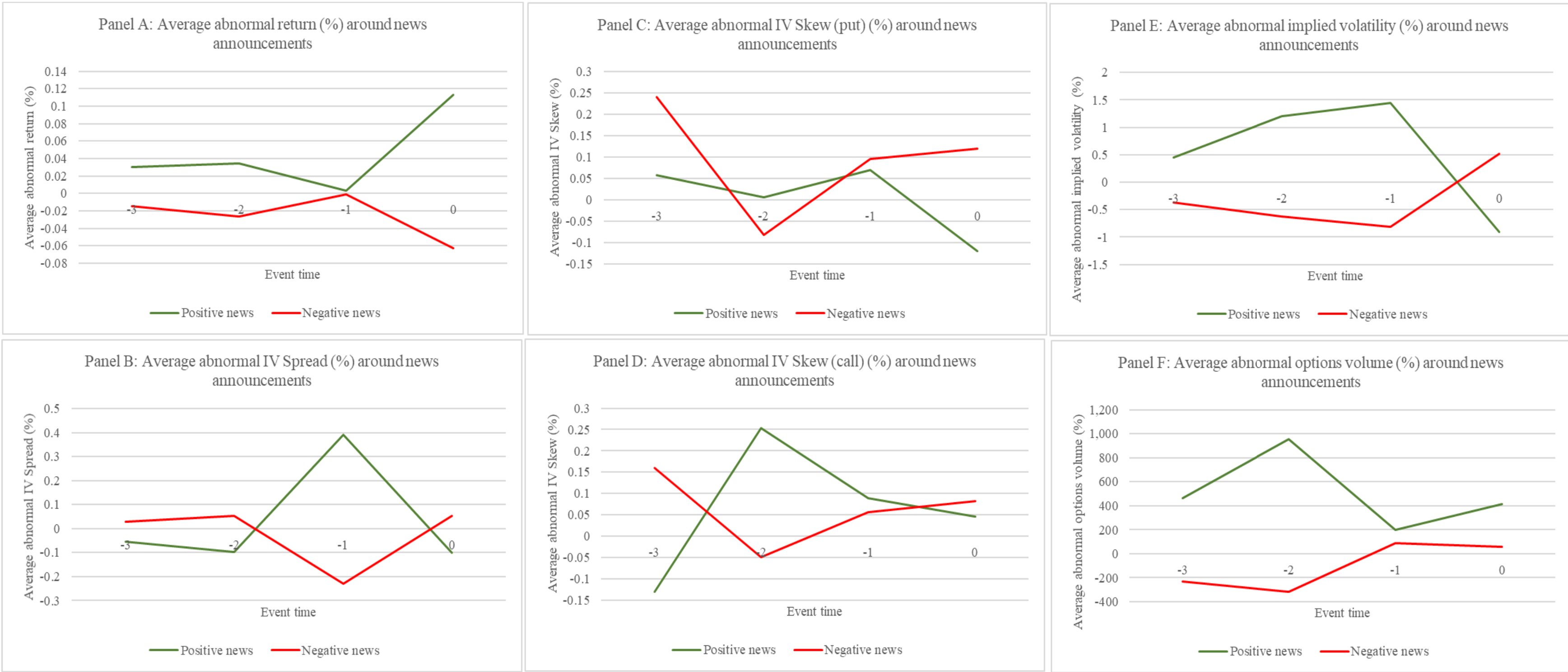


Figure 3.1: Abnormal metrics around surprise announcements

Silver

Table 3.1: News announcements

This table lists the macro-economic and commodity-specific news events used in the analysis between 2007 and 2017. News announcement dates were obtained from *Bloomberg*.

Macro-economic announcements relating to metal markets		
Event	N	Frequency
Consumer price index	134	Monthly
GDP announcement	33	Quarterly
Industrial production	134	Monthly
Commodity-specific announcements		
Event	N	Frequency
Short-term energy outlook (crude and natural gas)	134	Monthly
WASDE report (corn, soybeans and wheat)	134	Monthly

Table 3.2: Descriptive statistics of commodity derivatives

This table shows the daily descriptive statistics for futures and options in the sample between 2007 and 2017. N is the number of observations. Descriptive statistics are calculated from daily values.

Agriculture

	Mean	Median	Std dev.
Corn			
Number of options		25,805	
Futures return (%)	–0.100	–0.045	0.296
Futures volume	130,996.2	123,101.0	85,098.6
Options volume	35,523.6	23,627.0	41,112.2
Options open interest	8,554.6	8,134.9	6,141.5
Implied volatility (%)	31.962	31.437	92.188
Soybeans			
Number of options		40,414	
Futures return (%)	–0.025	0.054	0.224
Futures volume	70,554.6	55,556.0	56,738.6
Options volume	18,721.8	11,002.0	23,388.7
Options open interest	4,212.4	3,843.2	3,382.0
Implied volatility (%)	25.913	23.857	118.173
Wheat			
Number of options		21,251	
Futures return (%)	0.687	0.428	0.296
Futures volume	53,731.5	50,628.0	27,105.9
Options volume	8,942.5	7,174.0	9,699.3
Options open interest	2,450.5	2,486.9	1,545.2
Implied volatility (%)	35.450	35.184	96.884

Table 3.2: Descriptive statistics (continued)

Energy

	Mean	Median	Std dev.
Crude oil			
Number of options		402,287	
Futures return (%)	0.009	–0.065	0.234
Futures volume	144,382.5	92,467.0	136,112.1
Options volume	24,254.1	17,543.0	26,511.5
Options open interest	4,716.7	4,250.4	3,012.8
Implied volatility (%)	35.490	32.866	242.251
Natural gas			
Number of options		288,407	
Futures return (%)	0.359	0.287	0.312
Futures volume	71,980.5	54,366.0	58,713.8
Options volume	10,478.4	3,047.0	17,540.9
Options open interest	3,682.7	3,194.0	1,973.2
Implied volatility (%)	48.513	43.534	58.995

Metals

	Mean	Median	Std dev.
Gold			
Number of options		99,677	
Futures return (%)	0.029	–0.008	0.121
Futures volume	72,745.8	31,253	77,149.2
Options volume	7,922.5	5,203.0	9,395.2
Options open interest	2,037.2	1,993.0	1,091.1
Implied volatility (%)	21.200	19.641	142.974
Silver			
Number of options		89,427	
Futures return (%)	–0.021	–0.152	0.216
Futures volume	23,794.3	19,937	24,484.3
Options volume	1,639.6	1,133.0	1,954.3
Options open interest	698.2	622.4	453.3
Implied volatility (%)	35.150	33.142	162.172

Table 3.3: Implied volatility descriptive statistics

The table shows descriptive statistics for IV spread, IV skew (put) and IV skew (call). The sample includes futures and options data between 2007 and 2017.

Agriculture

	Mean	Median	Std dev.
Corn			
Number of options		25,805	
IV spread (%)	0.467	0.269	9.219
IV skew (put) (%)	2.255	1.696	3.900
IV skew (call) (%)	1.847	0.972	3.562
Soybeans			
Number of options		40,414	
IV spread (%)	0.383	0.157	11.817
IV skew (put) (%)	3.207	2.319	4.131
IV skew (call) (%)	2.204	1.471	3.582
Wheat			
Number of options		21,251	
IV spread (%)	0.174	0.364	9.688
IV skew (put) (%)	0.478	–0.105	3.246
IV skew (call) (%)	–0.321	–0.654	2.560

Energy

	Mean	Median	Std dev.
Crude oil			
Number of options		402,287	
IV spread (%)	–0.289	0.015	2.423
IV skew (put) (%)	6.149	5.202	4.424
IV skew (call) (%)	6.109	4.261	6.552
Natural gas			
Number of options		288,407	
IV spread (%)	0.336	0.103	5.899
IV skew (put) (%)	0.688	0.563	5.963
IV skew (call) (%)	0.191	0.313	7.247

Table 3.3: Implied volatility descriptive statistics (continued)

Metals

	Mean	Median	Std dev.
Gold			
Number of options		99,677	
IV spread (%)	0.489	0.019	1.430
IV skew (put) (%)	5.565	4.313	5.564
IV skew (call) (%)	5.401	3.668	5.773
Silver			
Number of options		89,427	
IV spread (%)	0.605	0.028	1.622
IV skew (put) (%)	4.444	3.012	5.132
IV skew (call) (%)	4.162	2.801	4.118

Table 3.4: Panel regression results

This table shows that the results of panel regressions of the form $METRIC_{i,t} = \beta_0 + \beta_1 EventDV + \beta_2 Controls + \varepsilon$. $METRIC_{i,t}$ is IV spread, IV skew (put) or IV skew (call). $EventDV$ is equal to 1 during the period [–3,–1]. The regression controls for event-fixed effects and control variables include lag return, lag options volume and lag options open interest (all variable definitions are provided in Appendix A). The sample includes futures and options data between 2007 and 2017. *t*-statistic values are shown in brackets. ***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent level.

Corn							
Panel A: IV spread							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	4.086***	3.754***	2.295**	*Intercept*	0.189	0.279	–0.158
	(6.27)	(2.98)	(2.13)		(0.44)	(0.67)	(–0.36)
EventDV	–0.207	–0.282	0.010	*EventDV*	–0.442	0.412	0.783*
	(–0.98)	(–1.08)	(0.04)		(–1.63)	(1.05)	(1.90)
Lag(Return)		–0.084***	–0.084***	*Lag(Return)*		–0.090**'	–0.089***
		(–9.35)	(–9.35)			(–9.35)	(–9.34)
Lag(Options Volume)			0.120***	*Lag(Options Volume)*			0.132***
			(5.78)				(6.18)
Lag(OptionsOI)			–0.013***	*Lag(OptionsOI)*			–0.014***
			(–3.29)				(–3.06)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	2,876	2,876	2,876	*N*	2,566	2,566	2,566
Adj. R2	0.020	0.098	0.083	*Adj. R2*	0.018	0.104	0.086
Panel B: IV skew (put)							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	–2.187***	–2.141***	0.387	*Intercept*	2.599***	2.592***	4.488***
	(–4.46)	(–4.22)	(0.54)		(10.50)	(10.49)	(16.84)
EventDV	1.225*	1.232*	0.945	*EventDV*	1.190***	1.088**	1.001**
	(1.74)	(1.75)	(1.52)		(2.03)	(1.85)	(1.82)
Lag(Return)		0.117***	0.108***	*Lag(Return)*		0.103***	0.094***
		(5.05)	(4.81)			(4.36)	(4.09)
Lag(Options Volume)			–0.018***	*Lag(Options Volume)*			–0.017***
			(–15.27)				(–14.09)
Lag(OptionsOI)			–0.119***	*Lag(OptionsOI)*			–0.135***
			(–6.87)				(–7.49)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	2,854	2,854	2,854	*N*	2,545	2,545	2,545
Adj. R2	0.079	0.086	0.205	*Adj. R2*	0.081	0.087	0.201
Panel C: IV skew (call)							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	–1.788***	–1.768***	0.7408	*Intercept*	2.201***	2.013***	3.683***
	(–4.42)	(–4.49)	(1.04)		(8.22)	(8.24)	(13.56)
EventDV	0.903	0.901	0.644	*EventDV*	0.644	0.609	0.599
	(1.42)	(1.42)	(1.11)		(1.06)	(1.00)	(1.07)
Lag(Return)		0.054***	0.040**	*Lag(Return)*		0.035*	0.023
		(2.75)	(2.07)			(1.75)	(1.16)
Lag(Options Volume)			–0.016***	*Lag(Options Volume)*			–0.016***
			(–14.08)				(–12.82)
Lag(OptionsOI)			–0.117***	*Lag(OptionsOI)*			–0.124***
			(–7.43)				(–7.45)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	2,570	2,570	2,570	*N*	2,308	2,308	2,308
Adj. R2	0.055	0.057	0.172	*Adj. R2*	0.056	0.057	0.167

Table 3.4: Panel regression results (continued)

Soybeans

Panel A: IV spread

Positive news				Negative news			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	0.372 (0.79)	0.452 (1.01)	–0.032 (–0.07)	*Intercept*	0.373 (0.81)	0.410 (0.94)	–0.056 (–0.13)
EventDV	–0.238 (–0.99)	–0.205 (–0.53)	0.017 (0.04)	*EventDV*	–0.335* (–1.72)	0.163 (0.57)	0.262 (0.92)
Lag(Return)		–0.195*** (–9.61)	–0.195*** (–9.62)	*Lag(Return)*		–0.173*** (–9.84)	–0.173*** (–9.86)
Lag(Options Volume)			0.210*** (5.90)	*Lag(Options Volume)*			0.189*** (6.92)
Lag(OptionsOI)			–0.015** (–2.53)	*Lag(OptionsOI)*			–0.013*** (–3.05)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	4,683	4,683	4,683	*N*	6,147	6,147	6,147
Adj. R2	0.003	0.138	0.109	*Adj. R2*	0.002	0.119	0.103

Panel B: IV skew (put)

Positive news				Negative news			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	3.754*** (18.13)	3.744*** (18.34)	5.032*** (23.73)	*Intercept*	3.670*** (17.89)	3.666*** (18.13)	4.963*** (24.26)
EventDV	0.715 (1.49)	0.715 (1.49)	0.586 (1.26)	*EventDV*	0.884** (2.07)	0.788* (1.86)	0.799** (2.01)
Lag(Return)		0.330*** (13.78)	0.327*** (14.01)	*Lag(Return)*		0.318*** (14.68)	0.317*** (14.99)
Lag(Options Volume)			–0.030*** (–14.98)	*Lag(Options Volume)*			–0.031*** (–18.43)
Lag(OptionsOI)			–0.034 (–1.16)	*Lag(OptionsOI)*			–0.017 (–0.70)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	4,655	4,655	4,655	*N*	6,114	6,114	6,114
Adj. R2	0.038	0.068	0.133	*Adj. R2*	0.035	0.062	0.132

Panel C: IV skew (call)

Positive news				Negative news			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	–1.644*** (–5.05)	–1.797*** (–4.71)	–1.068** (–2.34)	*Intercept*	2.591*** (13.16)	2.605*** (13.26)	3.647*** (18.46)
EventDV	0.510 (1.35)	0.479 (1.28)	0.589 (1.61)	*EventDV*	0.793* (1.85)	0.755* (1.76)	0.799** (2.01)
Lag(Return)		0.083*** (5.19)	0.087*** (5.52)	*Lag(Return)*		0.093*** (4.50)	0.095*** (4.65)
Lag(Options Volume)			–0.019*** (–5.41)	*Lag(Options Volume)*			–0.022*** (–14.61)
Lag(OptionsOI)			–0.351*** (–5.33)	*Lag(OptionsOI)*			–0.064*** (–2.65)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	2,211	2,211	2,211	*N*	5,228	5,228	5,228
Adj. R2	0.026	0.036	0.08	*Adj. R2*	0.031	0.034	0.096

Table 3.4: Panel regression results (continued)

Wheat							
Panel A: IV spread							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	2.851 (0.64)	4.445 (1.11)	3.333 (0.85)	*Intercept*	4.269* (1.65)	5.839*** (5.18)	3.840** (2.42)
EventDV	0.075 (0.30)	0.335 (0.89)	0.661 (1.57)	*EventDV*	–2.243 (–1.24)	–2.464 (–1.43)	–2.035 (–1.18)
Lag(Return)		–0.086*** (–4.81)	–0.087*** (–4.84)	*Lag(Return)*		–0.080*** (–5.82)	–0.081*** (–5.87)
Lag(Options Volume)			0.698*** (4.92)	*Lag(Options Volume)*			0.611*** (6.63)
Lag(OptionsOI)			–0.116*** (–4.77)	*Lag (OptionsOI)*			–0.086*** (–6.05)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	2,655	2,655	2,655	*N*	4,615	4,615	4,615
Adj. R2	0.011	0.076	0.096	*Adj. R2*	0.008	0.066	0.093
Panel B: IV skew (put)							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	–1.787** (–2.25)	–2.035*** (–2.66)	–0.328 (–0.33)	*Intercept*	–2.260*** (–5.5)	–2.531*** (–3.78)	0.468*** (3.01)
EventDV	1.268** (2.27)	1.227** (2.20)	1.377*** (2.68)	*EventDV*	0.563 (1.37)	0.601 (1.47)	0.93** (2.43)
Lag(Return)		0.134*** (6.96)	0.151*** (8.02)	*Lag(Return)*		0.138*** (9.35)	0.145*** (10.15)
Lag(Options Volume)			–0.054*** (–13.74)	*Lag(Options Volume)*			–0.054*** (–19.34)
Lag(OptionsOI)			–0.441*** (–7.92)	*Lag (OptionsOI)*			–0.402*** (–11.04)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	2,622	2,622	2,622	*N*	4,554	4,554	4,554
Adj. R2	0.045	0.060	0.172	*Adj. R2*	0.039	0.054	0.171
Panel C: IV skew (call)							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	–1.644*** (–5.05)	–1.797*** (–4.71)	–1.068** (–2.34)	*Intercept*	–1.644*** (–5.05)	–1.797*** (–4.71)	–1.068** (–2.34)
EventDV	0.510 (1.35)	0.479 (1.28)	0.589 (1.61)	*EventDV*	0.510 (1.35)	0.479 (1.28)	0.589 (1.61)
Lag(Return)		0.083*** (5.19)	0.087*** (5.52)	*Lag(Return)*		0.083*** (5.19)	0.087*** (5.52)
Lag(Options Volume)			–0.019*** (–5.41)	*Lag(Options Volume)*			–0.019*** (–5.41)
Lag(OptionsOI)			–0.351*** (–5.33)	*Lag (OptionsOI)*			–0.351*** (–5.33)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	2,211	2,211	2,211	*N*	2,211	2,211	2,211
Adj. R2	0.026	0.036	0.08	*Adj. R2*	0.026	0.036	0.080

Table 3.4: Panel regression results (continued)

Crude oil

Panel A: IV spread

	Positive news				**Negative news**		
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	2.814 (1.23)	2.969 (1.28)	2.788 (1.17)	*Intercept*	0.323 (1.47)	0.441*** (4.96)	0.369*** (3.52)
EventDV	1.642*** (2.83)	1.604*** (2.77)	1.122** (1.92)	*EventDV*	0.551* (1.73)	0.507* (1.61)	0.340 (1.11)
Lag(Return)		–0.021** (–2.08)	–0.022** (–2.21)	*Lag(Return)*		–0.025* (–1.83)	–0.026* (–1.90)
Lag(Options Volume)			0.016 (0.33)	*Lag(Options Volume)*			–0.002 (–0.03)
Lag(Options OI)			0.075*** (8.97)	*Lag(Options OI)*			0.072*** (4.57)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	12,330	12,330	12,330	*N*	7,279	7,279	7,279
Adj. R2	0.102	0.103	0.182	*Adj. R2*	0.088	0.09	0.14

Panel B: IV skew (put)

	Positive news				**Negative news**		
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	10.070*** (3.49)	10.094*** (3.51)	10.139*** (3.53)	*Intercept*	15.309*** (6.94)	15.336*** (6.85)	15.294*** (6.94)
EventDV	0.551* (1.92)	0.546* (1.90)	0.695*** (2.49)	*EventDV*	0.216 (0.91)	0.207 (0.87)	0.328 (1.43)
Lag(Return)		–0.032*** (–2.96)	–0.027** (–2.55)	*Lag(Return)*		–0.055*** (–3.45)	–0.039** (–2.43)
Lag(Options Volume)			0.001 (0.11)	*Lag(Options Volume)*			0.006*** (2.96)
Lag(Options OI)			–0.230*** (–10.68)	*Lag(Options OI)*			–0.470*** (–9.76)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	12,302	12,302	12,302	*N*	7,269	7,269	7,269
Adj. R2	0.248	0.248	0.269	*Adj. R2*	0.274	0.275	0.301

Panel C: IV skew (call)

	Positive news				**Negative news**		
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	118.607*** (2.89)	118.479*** (2.89)	118.425*** (2.91)	*Intercept*	17.843*** (9.65)	17.861*** (9.69)	17.764*** (9.61)
EventDV	0.542** (2.02)	0.543** (2.03)	0.703*** (2.71)	*EventDV*	0.224 (0.93)	0.221 (0.92)	0.380 (1.64)
Lag(Return)		0.017 (1.20)	0.026* (1.88)	*Lag(Return)*		–0.037* (–1.74)	–0.019 (–0.89)
Lag(Options Volume)			0.003* (1.94)	*Lag(Options Volume)*			0.010*** (4.44)
Lag(Options OI)			–0.240*** (–8.61)	*Lag(Options OI)*			–0.540*** (–9.76)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	11,893	11,893	11,893	*N*	6,954	6,954	6,954
Adj. R2	0.202	0.202	0.216	*Adj. R2*	0.289	0.289	0.312

Table 3.4: Panel regression results (continued)

Natural gas

Panel A: IV spread

Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	–0.043** (–2.35)	–0.062*** (–3.03)	–0.275*** (–6.36)	*Intercept*	–0.042** (–2.31)	–0.128*** (–4.62)	–0.343*** (–7.19)
EventDV	–0.367*** (–2.66)	–0.347*** (–2.52)	–0.379*** (–2.76)	*EventDV*	–0.418*** (–3.19)	–0.425*** (–3.24)	–0.453*** (–3.45)
Lag(Return)		0.009*** (21.00)	0.009*** (21.17)	*Lag(Return)*		0.009*** (20.76)	0.009*** (20.93)
Lag(Options Volume)			0.023*** (2.94)	*Lag(Options Volume)*			0.019*** (2.77)
Lag(OptionsOI)			0.004*** (7.50)	*Lag(OptionsOI)*			0.003*** (6.60)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	139,364	139,364	139,364	*N*	148,931	148,931	148,931
Adj. R2	0.016	0.018	0.060	*Adj. R2*	0.015	0.017	0.057

Panel B: IV skew (put)

Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	1.863*** (11.64)	1.868*** (11.67)	2.066*** (12.71)	*Intercept*	1.844*** (11.57)	0.679*** (6.74)	1.188*** (10.14)
EventDV	0.811*** (6.93)	0.806*** (6.89)	0.954*** (8.27)	*EventDV*	0.229 (1.49)	0.233* (1.51)	0.425*** (2.77)
Lag(Return)		–0.025*** (–3.93)	–0.033*** (–5.24)	*Lag(Return)*		–0.022*** (–3.65)	–0.031*** (–5.07)
Lag(Options Volume)			0.028*** (23.75)	*Lag(Options Volume)*			0.028*** (25.24)
Lag(OptionsOI)			–0.525*** (–61.24)	*Lag(OptionsOI)*			–0.530*** (–65.17)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	139,157	139,157	139,157	*N*	148,714	148,714	148,714
Adj. R2	0.048	0.048	0.060	*Adj. R2*	0.048	0.048	0.061

Panel C: IV skew (call)

Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	1.184*** (8.13)	1.192*** (8.19)	0.698*** (4.44)	*Intercept*	11.689*** (8.07)	3.174*** (3.50)	0.148 (0.13)
EventDV	0.678*** (5.38)	0.664*** (5.28)	0.730*** (5.85)	*EventDV*	0.123 (0.61)	0.132 (0.66)	0.234 (1.17)
Lag(Return)		–0.064*** (–8.20)	–0.073*** (–9.38)	*Lag(Return)*		–0.062*** (–8.27)	–0.071*** (–9.54)
Lag(Options Volume)			0.041*** (23.48)	*Lag(Options Volume)*			0.040*** (24.71)
Lag(OptionsOI)			–0.491*** (–45.69)	*Lag(OptionsOI)*			–0.490*** (–48.38)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	135,196	135,196	135,196	*N*	144,410	144,410	144,410
Adj. R2	0.026	0.027	0.036	*Adj. R2*	0.027	0.027	0.036

Table 3.4: Panel regression results (continued)

Gold

Panel A: IV spread

Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	0.192***	0.159***	–5.543***	*Intercept*	0.197***	0.216***	–6.368***
	(6.08)	(4.38)	(–19.34)		(6.09)	(5.16)	(–12.88)
EventDV	–3.670***	–3.673***	–2.262***	*EventDV*	–4.098***	–4.100***	–2.651***
	(–6.52)	(–6.52)	(–4.23)		(–11.10)	(–11.05)	(–7.27)
Lag(Return)		0.035***	0.035***	*Lag(Return)*		0.031***	0.033***
		(6.30)	(6.62)			(7.51)	(8.61)
Lag(Options Volume)			5.080***	*Lag(Options Volume)*			5.171***
			(42.89)				(58.09)
Lag(Options OI)			–0.330***	*Lag(Options OI)*			–0.368***
			(–31.99)				(–44.97)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	35,436	35,436	35,436	*N*	64,241	64,241	64,241
Adj. R2	0.129	0.130	0.124	*Adj. R2*	0.114	0.115	0.1238

Panel B: IV skew (put)

Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	7.486***	7.466***	9.212***	*Intercept*	7.43***	3.794***	6.037***
	(28.79)	(28.75)	(37.06)		(28.74)	(34.73)	(46.69)
EventDV	1.024***	1.019***	1.056***	*EventDV*	1.181***	1.174***	1.136***
	(3.97)	(3.95)	(4.25)		(6.50)	(6.48)	(6.43)
Lag(Return)		0.214***	0.254***	*Lag(Return)*		0.235***	0.269***
		(10.25)	(11.99)			(14.98)	(16.92)
Lag(Options Volume)			–0.076***	*Lag(Options Volume)*			–0.081***
			(–29.02)				(–41.17)
Lag(Options OI)			–0.200***	*Lag(Options OI)*			–0.044**
			(–7.64)				(–2.13)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	35,391	35,391	35,391	*N*	64,168	64,168	64,168
Adj. R2	0.099	0.101	0.127	*Adj. R2*	0.087	0.089	0.115

Panel C: IV skew (call)

Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	54.647***	54.379***	54.554***	*Intercept*	54.255***	28.620***	31.313***
	(25.6)	(25.56)	(24.71)		(25.58)	(19.96)	(17.86)
EventDV	–0.505***	–0.520***	–0.178**	*EventDV*	–0.123	–0.126	0.173
	(–2.09)	(–2.16)	(–1.74)		(–0.63)	(–0.64)	(0.88)
Lag(Return)		0.304***	0.337***	*Lag(Return)*		0.319***	0.357***
		(14.54)	(16.39)			(20.55)	(23.33)
Lag(Options Volume)			0.062***	*Lag(Options Volume)*			0.058***
			(15.08)				(18.98)
Lag(Options OI)			–0.955***	*Lag(Options OI)*			–0.928***
			(–26.34)				(–32.66)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	33,406	33,406	33,406	*N*	60,681	60,681	60,681
Adj. R2	0.008	0.012	0.033	*Adj. R2*	0.007	0.012	0.031

Table 3.4: Panel regression results (continued)

Silver

Panel A: IV spread

Positive news				Negative news			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	0.347*** (4.56)	0.312*** (3.52)	–5.936*** (–18.73)	*Intercept*	0.318*** (4.43)	0.214*** (3.34)	–9.553*** (–19.97)
EventDV	–4.636*** (–8.15)	–4.683*** (–8.29)	–3.883*** (–6.82)	*EventDV*	–3.190*** (–6.32)	–3.154*** (–6.23)	–1.968*** (–4.10)
Lag(Return)		0.050*** (8.87)	0.043*** (8.28)	*Lag(Return)*		0.048*** (11.56)	0.041*** (10.74)
Lag(Options Volume)			11.870*** (45.87)	*Lag(Options Volume)*			12.299*** (62.74)
Lag(Options OI)			–0.587*** (–18.18)	*Lag(Options OI)*			–0.691*** (–27.96)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	31,792	31,792	31,792	*N*	57,635	57,635	57,635
Adj. R2	0.110	0.114	0.145	*Adj. R2*	0.099	0.103	0.144

Panel B: IV skew (put)

Positive news				Negative news			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	7.19*** (24.49)	7.186*** (24.50)	8.925*** (31.74)	*Intercept*	7.157*** (24.58)	3.387*** (22.25)	5.932*** (30.13)
EventDV	1.177*** (4.99)	1.171*** (4.97)	1.035*** (4.62)	*EventDV*	1.449*** (7.71)	1.460*** (7.77)	1.211*** (6.71)
Lag(Return)		0.064*** (6.03)	0.081*** (7.5)	*Lag(Return)*		0.068*** (8.39)	0.085*** (10.38)
Lag(Options Volume)			–0.280*** (–63.15)	*Lag(Options Volume)*			–0.295*** (–87.88)
Lag(Options OI)			0.362*** (2.91)	*Lag(Options OI)*			1.073*** (11.43)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	31,684	31,684	31,684	*N*	57,441	57,441	57,441
Adj. R2	0.199	0.200	0.257	*Adj. R2*	0.177	0.178	0.241

Panel C: IV skew (call)

Positive news				Negative news			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	4.935*** (19.22)	4.935*** (19.23)	5.555*** (21.97)	*Intercept*	4.917*** (19.29)	1.747*** (12.95)	2.663*** (18.69)
EventDV	–0.042 (–0.21)	–0.054 (–0.26)	–0.065 (–0.32)	*EventDV*	0.471*** (2.83)	0.481*** (2.89)	0.445*** (2.70)
Lag(Return)		0.068*** (6.41)	0.064*** (5.90)	*Lag(Return)*		0.072*** (9.03)	0.069*** (8.37)
Lag(Options Volume)			–0.081*** (–16.35)	*Lag(Options Volume)*			–0.086*** (–22.79)
Lag(Options OI)			–0.426*** (–4.12)	*Lag(OptionsOI)*			–0.17** (–2.15)
Event-fixed effects	Yes	Yes	Yes	*Event-fixed effects*	Yes	Yes	Yes
N	28,911	28,911	28,911	*N*	52,533	52,533	52,533
Adj. R2	0.033	0.034	0.042	*Adj. R2*	0.027	0.029	0.037

Table 3.5: Return-based panel regression results

This table shows the results of panel regressions of the form $RET_{i,t} = \beta_0 + \beta_1 AbnMETRIC[-3,-1] + \beta_2 Controls + \varepsilon$. $RET_{i,t}$ is abnormal announcement date futures return. $AbnMETRIC[-3,-1]$ is the abnormal average IV spread, IV skew (put) or IV skew (put). Abnormal IV spread or IV skew is calculated as the difference between the actual value on day *t* and the average value during the annual benchmark period. Control variables include abnormal pre-announcement return (*Abn. Return (%) [–3,–1]), options volume* and *options open interest* (all variable definitions are provided in Appendix A). Year dummies are included. The sample includes futures and options data between 2007 and 2017. *t*-statistic values are shown in brackets. ***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent level.

Corn

Panel A: Abn. IV spread

Variable	Positive news: Model 1	Model 2	Model 3	Variable	Negative news: Model 1	Model 2	Model 3
Intercept	2.591*** (2.74)	2.309** (2.53)	2.457** (2.03)	*Intercept*	–3.059*** (–3.58)	–2.911*** (–3.24)	–3.664*** (–2.88)
Abn. IV spread [–3,–1]	0.137*** (2.35)	0.101** (1.81)	0.103** (1.85)	*Abn. IV spread [–3,–1]*	0.989 (0.19)	0.561 (0.11)	0.285 (0.06)
Abn. return (%) [–3,–1]		–0.332*** (–4.00)	–0.333*** (–3.98)	*Abn. return (%) [–3,–1]*		–0.090 (–0.75)	–0.097 (–0.83)
Options volume			0.061 (0.61)	*Options volume*			–0.050 (–0.50)
Options open interest			0.020 (0.08)	*Options open interest*			0.255 (1.02)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	160	160	160	*N*	159	159	159
Adj. R2	0.131	0.206	0.209	*Adj. R2*	0.077	0.082	0.088

Panel B: Abn. IV skew (put)

Variable	Positive news: Model 1	Model 2	Model 3	Variable	Negative news: Model 1	Model 2	Model 3
Intercept	3.208*** (3.25)	2.714*** (2.89)	2.770** (2.19)	*Intercept*	–2.893*** (–3.42)	–2.769*** (–3.14)	–3.530*** (–2.76)
Abn. IV skew (put) [–3,–1]	–0.009 (–0.10)	0.009 (0.11)	0.018 (0.21)	*Abn. IV skew (put) [–3,–1]*	0.010 (1.20)	0.011 (1.23)	0.013* (1.52)
Abn. return (%) [–3,–1]		–0.372*** (–4.26)	–0.374*** (–4.18)	*Abn. return (%) [–3,–1]*		–0.088 (–0.70)	–0.099 (–0.79)
Options volume			0.052 (0.52)	*Options volume*			–0.021 (–0.21)
Options open interest			0.045 (0.18)	*Options open interest*			0.305 (1.22)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	160	160	160	*N*	159	159	159
Adj. R2	0.083	0.181	0.183	*Adj. R2*	0.081	0.086	0.094

Panel C: Abn. IV skew (call)

Variable	Positive news: Model 1	Model 2	Model 3	Variable	Negative news: Model 1	Model 2	Model 3
Intercept	3.148*** (3.07)	2.499** (2.53)	2.607* (1.88)	*Intercept*	–2.271*** (–2.65)	–1.938** (–2.26)	–3.253** (–2.55)
Abn. IV skew (call) [–3,–1]	0.006 (0.61)	0.010 (1.17)	0.011 (1.19)	*Abn. IV skew (call) [–3–1]*	0.019* (1.78)	0.021** (1.94)	0.027*** (2.56)
Abn. return (%) [–3,–1]		–0.409*** (–3.77)	–0.412*** (–3.74)	*Abn. Return (%) [–3,–1]*		–0.198 (–1.40)	–0.254* (–1.83)
Options volume			0.060 (0.60)	*Options volume*			–0.055 (–0.55)
Options open interest			0.053 (0.21)	*Options open interest*			0.523** (2.09)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	160	160	160	*N*	159	159	159
Adj. R2	0.085	0.178	0.182	*Adj. R2*	0.071	0.094	0.122

Table 3.5: Return-based panel regression results (continued)

Soybeans							
Panel A: Abn. IV spread							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	3.137*** (5.06)	2.676*** (3.53)	1.762* (1.83)	*Intercept*	–2.620*** (–3.55)	–2.469*** (–3.05)	–1.456 (–1.53)
Abn. IV spread [–3,–1]	0.338 (1.23)	0.212 (0.67)	0.230 (0.72)	*Abn. IV spread [–3,–1]*	0.957*** (3.81)	0.855*** (3.97)	0.903*** (4.34)
Abn. return (%) [–3,–1]		–0.356** (–2.36)	–0.358** (–2.39)	*Abn. return (%) [–3,–1]*		–0.090 (–0.82)	–0.073 (–0.65)
Options volume			–0.303*** (–3.03)	*Options volume*			0.287*** (2.87)
Options open interest			0.308 (1.23)	*Options open interest*			–0.798*** (–3.19)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	234	234	234	*N*	206	206	206
Adj. R2	0.148	0.224	0.246	*Adj. R2*	0.170	0.175	0.206
Panel B: Abn. IV skew (put)							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	3.067*** (4.73)	2.626*** (3.47)	1.732* (1.76)	*Intercept*	–2.889*** (–3.56)	–2.431*** (–2.87)	–1.472 (–1.47)
Abn. IV skew (put) [–3,–1]	–0.003 (–0.62)	–0.001 (–0.15)	–0.001 (–0.26)	*Abn. IV skew (put) [–3,–1]*	0.001 (0.17)	0.002 (0.47)	0.002 (0.47)
Abn. return (%) [–3,–1]		–0.369** (–2.43)	–0.372** (–2.45)	*Abn. return (%) [–3,–1]*		–0.211* (–1.85)	–0.202* (–1.73)
Options volume			–0.292*** (–2.92)	*Options volume*			0.271*** (2.71)
Options open interest			0.273 (1.09)	*Options open interest*			–0.655*** (–2.62)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	234	234	234	*N*	206	206	206
Adj. R2	0.137	0.217	0.238	*Adj. R2*	0.101	0.128	0.155
Panel C: Abn. IV skew (call)							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	2.807*** (4.41)	1.993*** (3.00)	0.924 (1.06)	*Intercept*	–2.326*** (–2.72)	–1.888** (–2.08)	–0.862 (–0.81)
Abn. IV skew (call) [–3,–1]	0.516 (1.03)	0.744 (1.51)	0.640 (1.25)	*Abn. IV skew (call) [–3,–1]*	1.036** (1.95)	1.126*** (2.21)	1.061** (2.11)
Abn. return (%) [–3,–1]		–0.517*** (–3.80)	–0.521*** (–3.92)	*Abn. return (%) [–3,–1]*		–0.203 (–1.55)	–0.201 (–1.51)
Options volume			–0.311*** (–3.11)	*Options volume*			0.250** (2.50)
Options open Interest			0.398 (1.59)	*Options open interest*			–0.62** (–2.48)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	234	234	234	*N*	206	206	206
Adj. R2	0.115	0.262	0.290	*Adj. R2*	0.098	0.124	0.152

Table 3.5: Return-based panel regression results (continued)

Wheat

Panel A: Abn. IV spread

Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	4.250*** (4.01)	3.776*** (4.02)	3.621*** (2.72)	*Intercept*	–3.092*** (–3.11)	–2.808*** (–2.74)	–2.943** (–2.25)
Abn. IV spread [–3,–1]	0.206*** (10.66)	0.211*** (13.54)	0.210*** (13.03)	*Abn. IV spread [–3,–1]*	0.698 (1.32)	0.519 (0.97)	0.510 (0.96)
Abn. return (%) [–3,–1]		–0.477*** (–3.63)	–0.482*** (–3.63)	*Abn. return (%) [–3,–1]*		–0.151 (–1.32)	–0.150 (–1.31)
Options volume			–0.014 (–0.14)	*Options volume*			–0.014 (–0.14)
Options open interest			–0.098 (–0.39)	*Options open interest*			0.153 (0.61)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	128	128	128	*N*	131	131	131
Adj. R2	0.309	0.446	0.447	*Adj. R2*	0.107	0.123	0.125

Panel B: Abn. IV skew (put)

Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	5.116*** (3.77)	4.845*** (3.94)	5.117*** (3.00)	*Intercept*	–2.675*** (–2.70)	–2.389** (–2.30)	–2.695* (–1.96)
Abn. IV skew (put) [–3,–1]	–0.005 (–0.56)	–0.010 (–1.04)	–0.011 (–1.08)	*Abn. IV skew (put) [–3,–1]*	–0.012 (–0.81)	–0.010 (–0.69)	–0.008 (–0.48)
Abn. return (%) [–3,–1]		–0.480*** (–3.80)	–0.494*** (–3.88)	*Abn. return (%) [–3,–1]*		–0.190 (–1.46)	–0.190 (–1.45)
Options volume			0.047 (0.47)	*Options volume*			–0.025 (–0.25)
Options open interest			–0.243 (–0.97)	*Options open interest*			0.098 (0.39)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	128	128	128	*N*	131	131	131
Adj. R2	0.168	0.306	0.312	*Adj. R2*	0.084	0.112	0.113

Panel C: Abn. IV skew (call)

Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	5.807*** (4.08)	4.903*** (3.81)	5.445*** (3.17)	*Intercept*	–3.045*** (–2.93)	–2.485** (–2.25)	–2.825* (–1.72)
Abn. IV skew (call) [–3,–1]	0.233 (0.20)	0.494 (0.47)	0.298 (0.24)	*Abn. IV skew (call) [–3,–1]*	0.952 (0.58)	0.433 (0.27)	0.391 (0.22)
Abn. return (%) [–3,–1]		–0.535*** (–3.12)	–0.539*** (–3.15)	*Abn. return (%) [–3,–1]*		–0.191 (–1.21)	–0.192 (–1.23)
Options volume			0.075 (0.75)	*Options volume*			–0.028 (–0.28)
Options open interest			–0.185 (–0.74)	*Options open interest*			0.220 (0.88)
Year Dummies	Yes	Yes	Yes	*Year Dummies*	Yes	Yes	Yes
N	128	128	128	*N*	131	131	131
Adj. R2	0.230	0.357	0.363	*Adj. R2*	0.060	0.084	0.090

Table 3.5: Return-based panel regression results (continued)

Crude oil							
Panel A: Abn. IV spread							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	1.194*** (3.31)	1.193*** (3.31)	1.324** (2.47)	*Intercept*	–1.372*** (–3.60)	–1.324*** (–3.62)	–1.365** (–2.55)
Abn. IV spread [–3,–1]	0.376* (1.71)	0.397* (1.88)	0.328* (1.72)	*Abn. IV spread [–3,–1]*	–0.192 (–0.79)	–0.070 (–0.28)	–0.033 (–0.13)
Abn. return (%) [–3,–1]		–0.039 (–0.77)	–0.036 (–0.72)	*Abn. return (%) [–3,–1]*		–0.280*** (–3.53)	–0.275*** (–3.52)
Options volume			–0.057 (–0.57)	*Options volume*			0.070 (0.70)
Options open interest			–0.853*** (–3.41)	*Options open interest*			0.658*** (2.63)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	1,071	1,071	1,071	*N*	1,137	1,137	1,137
Adj. R2	0.015	0.016	0.035	*Adj. R2*	0.016	0.056	0.069
Panel B: Abn. IV skew (put)							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	1.173*** (3.24)	1.171*** (3.22)	1.214** (2.26)	*Intercept*	–1.341*** (–3.54)	–1.298*** (–3.56)	–1.278** (–2.41)
Abn. IV skew (put) [–3,–1]	0.027 (1.34)	0.027 (1.34)	0.030 (1.47)	*Abn. IV skew (put) [–3,–1]*	–0.005*** (–2.72)	–0.005** (–2.57)	–0.005*** (–2.80)
Abn. return (%) [–3,–1]		–0.032 (–0.63)	–0.030 (–0.60)	*Abn. return (%) [–3,–1]*		–0.278*** (–3.53)	–0.272*** (–3.50)
Options volume			–0.097 (–0.97)	*Options volume*			0.094 (0.94)
Options open interest			–0.810*** (–3.24)	*Options open interest*			0.635** (2.54)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	1,071	1,071	1,071	*N*	1,137	1,137	1,137
Adj. R2	0.013	0.014	0.034	*Adj. R2*	0.020	0.059	0.073
Panel C: Abn. IV skew (call)							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	1.270*** (3.52)	1.270*** (3.52)	1.449*** (2.73)	*Intercept*	–1.378*** (–3.63)	–1.348*** (–3.67)	–1.369** (–2.53)
Abn. IV skew (call) [–3,–1]	–0.081 (–0.75)	–0.072 (–0.67)	–0.035 (–0.31)	*Abn. IV skew (call) [–3,–1]*	–0.144* (–1.65)	–0.111 (–1.31)	–0.134 (–1.51)
Abn. return (%) [–3,–1]		–0.060 (–1.19)	–0.059 (–1.18)	*Abn. return (%) [–3,–1]*		–0.262*** (–3.25)	–0.257*** (–3.21)
Options volume			–0.040 (–0.40)	*Options volume*			0.067 (0.67)
Options open interest			–0.915*** (–3.66)	*Options open interest*			0.603** (2.41)
Year Dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	1,071	1,071	1,071	*N*	1,137	1,137	1,137
Adj. R2	0.015	0.017	0.037	*Adj. R2*	0.017	0.051	0.063

Table 3.5: Return-based panel regression results (continued)

Natural gas							
Panel A: Abn. IV spread							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	2.839*** (7.55)	2.755*** (7.68)	0.79 (1.05)	*Intercept*	–2.408*** (–5.25)	–2.425*** (–5.35)	0.778 (0.90)
Abn. IV spread [–3,–1]	0.918*** (4.67)	0.767*** (4.48)	0.743*** (4.69)	*Abn. IV spread [–3,–1]*	–0.148 (–0.65)	–0.182 (–0.78)	–0.112 (–0.49)
Abn. return (%) [–3,–1]		–0.256*** (–6.73)	–0.252*** (–6.68)	*Abn. return (%) [–3,–1]*		–0.216*** (–3.43)	–0.214*** (–3.38)
Options volume			–0.524*** (–5.24)	*Options volume*			0.594*** (5.94)
Options open interest			0.523** (2.09)	*Options open interest*			–0.433* (–1.73)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	849	849	849	*N*	784	784	784
Adj. R2	0.078	0.136	0.171	*Adj. R2*	0.031	0.054	0.095
Panel B: Abn. IV skew (put)							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	2.866*** (7.61)	2.774*** (7.72)	0.779 (1.03)	*Intercept*	–2.433*** (–5.47)	–2.447*** (–5.57)	0.594 (0.68)
Abn. IV skew (put) [–3,–1]	0.052 (1.96)	0.042* (1.72)	0.040* (1.65)	*Abn. IV skew (put) [–3,–1]*	–0.120*** (–4.20)	–0.117*** (–3.91)	–0.113*** (–3.72)
Abn. return (%) [–3,–1]		–0.260*** (–6.77)	–0.256*** (–6.73)	*Abn. return (%) [–3,–1]*		–0.203*** (–3.20)	–0.201*** (–3.17)
Options volume			–0.519*** (–5.19)	*Options volume*			0.559*** (5.59)
Options open interest			0.510** (2.04)	*Options open interest*			–0.420* (–1.68)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	849	849	849	*N*	784	784	784
Adj. R2	0.071	0.131	0.166	*Adj. R2*	0.086	0.106	0.143
Panel C: Abn. IV skew (call)							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	2.847*** (7.54)	2.770*** (7.69)	0.643 (0.83)	*Intercept*	–2.38*** (–5.34)	–2.393*** (–5.44)	0.615 (0.71)
Abn. IV skew (call) [–3,–1]	0.592** (2.55)	0.533** (2.44)	0.517*** (2.35)	*Abn. IV skew (call) [–3,–1]*	–0.783*** (–2.67)	–0.752** (–2.45)	–0.726*** (–2.40)
Abn. return (%) [–3,–1]		–0.261*** (–6.84)	–0.255*** (–6.77)	*Abn. return (%) [–3,–1]*		–0.201*** (–3.14)	–0.201*** (–3.13)
Options volume			–0.525*** (–5.25)	*Options volume*			0.570*** (5.70)
Options open interest			0.498** (1.99)	*Options open interest*			–0.468* (–1.87)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	849	849	849	*N*	784	784	784
Adj. R2	0.075	0.135	0.172	*Adj. R2*	0.074	0.094	0.132

Table 3.5: Return-based panel regression results (continued)

Gold							
Panel A: Abn. IV spread							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	1.627***	1.566***	1.358***	*Intercept*	–1.138***	–1.123***	–0.973***
	(5.29)	(5.42)	(3.53)		(–4.76)	(–4.71)	(–2.89)
Abn. IV spread [–3,–1]	0.196	0.281***	0.344***	*Abn. IV spread [–3,–1]*	–0.117	–0.131	–0.236***
	(1.43)	(2.14)	(2.53)		(–1.50)	(–1.57)	(–2.77)
Abn. return (%) [–3,–1]		–0.352***	–0.353***	*Abn. return (%) [–3,–1]*		–0.111	–0.087
		(–3.99)	(–3.96)			(–1.56)	(–1.23)
Options volume			–0.121	*Options volume*			0.166*
			(–1.21)				(1.66)
Options open interest			–0.133	*Options open interest*			0.530**
			(–0.53)				(2.12)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	310	310	310	*N*	340	340	340
Adj. R2	0.092	0.161	0.167	*Adj. R2*	0.079	0.090	0.118
Panel B: Abn. IV skew (_put)							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	1.554***	1.454***	1.170***	*Intercept*	–1.067***	–1.046***	–0.810**
	(5.23)	(5.19)	(3.17)		(–4.57)	(–4.47)	(–2.5)
Abn. IV skew (put) [–3,–1]	0.024	0.022	0.023	*Abn. IV skew (put) [–3,–1]*	0.011	0.012	0.017
	(1.42)	(1.36)	(1.41)		(0.91)	(0.98)	(1.42)
Abn. return (%) [–3,–1]		–0.330***	–0.330***	*Abn. return (%) [–3,–1]*		–0.107	–0.084
		(–3.53)	(–3.52)			(–1.50)	(–1.17)
Options volume			–0.143	*Options volume*			0.173*
			(–1.43)				(1.73)
Options open interest			0.038	*Options open interest*			0.428*
			(0.15)				(1.71)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	310	310	310	*N*	340	340	340
Adj. R2	0.093	0.155	0.159	*Adj. R2*	0.077	0.088	0.110
Panel C: Abn. IV skew (call)							
Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	1.713***	1.599***	1.289***	*Intercept*	–1.080***	–1.062***	–0.891***
	(5.35)	(5.3)	(3.31)		(–4.22)	(–4.16)	(–2.60)
Abn. IV skew (call) [–3,–1]	0.450***	0.430***	0.455***	*Abn. IV skew (call) [–3,–1]*	0.024	0.029	0.032
	(2.18)	(2.20)	(2.35)		(0.17)	(0.20)	(0.22)
Abn. return (%) [–3,–1]		–0.324***	–0.325***	*Abn. return (%) [–3,–1]*		–0.105	–0.085
		(–3.33)	(–3.34)			(–1.45)	(–1.18)
Options volume			–0.160	*Options volume*			0.147
			(–1.60)				(1.47)
Options open interest			0.015	*Options open interest*			0.408
			(0.06)				(1.63)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	310	310	310	*N*	340	340	340
Adj. R2	0.106	0.162	0.168	*Adj. R2*	0.076	0.086	0.106

Table 3.5: Return-based panel regression results (continued)

Silver

Panel A: Abn. IV spread

Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	3.369*** (6.02)	2.853*** (5.30)	2.639*** (4.23)	*Intercept*	–2.23*** (–5.88)	–2.189*** (–5.74)	–2.049*** (–4.3)
Abn. IV spread [–3,–1]	–0.105 (–0.43)	–0.226 (–1.04)	–0.206 (–0.97)	*Abn. IV spread [–3,–1]*	0.307*** (2.35)	0.290*** (2.28)	0.281*** (2.18)
Abn. return (%) [–3,–1]		–0.350*** (–4.44)	–0.351*** (–4.41)	*Abn. return (%) [–3,–1]*		–0.127* (–1.72)	–0.121* (–1.68)
Options volume			–0.102 (–1.02)	*Options volume*			0.111 (1.11)
Options open interest			–0.058 (–0.23)	*Options open interest*			0.180 (0.72)
Year dummies	Yes	Yes	Yes	*Year Dummies*	Yes	Yes	Yes
N	273	273	273	*N*	311	311	311
Adj. R2	0.134	0.189	0.191	*Adj. R2*	0.112	0.126	0.131

Panel B: Abn. IV skew (put)

Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	3.380*** (5.98)	2.869*** (5.17)	2.616*** (4.06)	*Intercept*	–2.076*** (–5.61)	–2.041*** (–5.48)	–1.908*** (–4.06)
Abn. IV skew (put) [–3,–1]	0.001 (0.38)	0.003 (0.95)	0.004 (0.97)	*Abn. IV skew (put) [–3,–1]*	–0.004 (–1.29)	–0.003 (–1.24)	–0.003 (–1.19)
Abn. return (%) [–3,–1]		–0.345*** (–4.16)	–0.348*** (–4.17)	*Abn. return (%) [–3,–1]*		–0.133* (–1.76)	–0.126* (–1.73)
Options volume			–0.120 (–1.20)	*Options volume*			0.114 (1.14)
Options open interest			–0.085 (–0.34)	*Options open interest*			0.228 (0.91)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	273	273	273	*N*	311	311	311
Adj. R2	0.134	0.188	0.191	*Adj. R2*	0.098	0.113	0.119

Panel C: Abn. IV skew (call)

Positive news				**Negative news**			
Variable	Model 1	Model 2	Model 3	Variable	Model 1	Model 2	Model 3
Intercept	3.629*** (6.19)	3.135*** (5.41)	2.996*** (4.49)	*Intercept*	–2.056*** (–5.10)	–2.047*** (–5.07)	–1.947*** (–3.78)
Abn. IV spread [–3,–1]	0.057 (1.10)	0.067 (1.34)	0.066 (1.33)	*Abn. IV spread [–3,–1]*	–0.205 (–0.53)	–0.217 (–0.57)	–0.209 (–0.55)
Abn. return (%) [–3,–1]		–0.360*** (–4.18)	–0.358*** (–4.13)	*Abn. return (%) [–3,–1]*		–0.156* (–1.93)	–0.151* (–1.93)
Options volume			–0.088 (–0.88)	*Options volume*			0.114 (1.14)
Options open interest			–0.120 (–0.48)	*Options open interest*			0.265 (1.06)
Year dummies	Yes	Yes	Yes	*Year dummies*	Yes	Yes	Yes
N	273	273	273	*N*	311	311	311
Adj. R2	0.153	0.209	0.212	*Adj. R2*	0.083	0.104	0.112

Appendix A – Chapter 3

Definition of variables

Variable	Definition
Abn. IV Spread [X,Y]	Abnormal IV spread averaged over days [X,Y]. Eq. (3.2) specifies the calculation of IV spread.
Abn. IV Skew (call) [X,Y]	Abnormal IV skew (call) averaged over days [X,Y]. Eq. (3.4) specifies the calculation of IV skew (call).
Abn. IV Skew (put) [X,Y]	Abnormal IV skew (put) averaged over days [X,Y]. Eq. (3.3) specifies the calculation of IV skew (put).
Abn. Return (%)	Abnormal daily return of the underlying futures contract is calculated as daily percentage return minus the average return of the benchmark period defined in Section 3.4.
EventDV	Dummy variable that equals one if the day is in the event window and zero if it is in the benchmark window.
Lag(OptionsVolume)	One–day lagged number of options contracts traded per trading day.
IV Skew (call) [X,Y]	IV skew (call) is calculated from options with the same time to maturity as the average IV across OTM call options minus the average across ATM call options (where moneyness is defined in Eq. [3.1]) per trading day.
IV Skew (put) [X,Y]	IV skew (put) is calculated from options with the same time to maturity as the average IV across OTM put options minus the average across ATM call options (in which moneyness is defined in Eq. [3.3]) per trading day.
IV Spread [X,Y]	IV spread is the difference in the IV between one pair of call and put options on the same underlying futures contract with the same strike price and expiration date. The IV difference is averaged across all the call and put pairs per trading day and is weighted by open interest. Eq. (3.2) outlines the definition of IV spread.
Lag(Return)	One–day lagged return of the underlying futures contract.
Options Open Interest	Number of contracts outstanding in each option at any trading day.
Options Volume	Number of options contracts traded per trading day.

Chapter 4: Price Discovery in Commodity Derivatives: Speculation or Hedging?

4.1 Research Question

This chapter answers the following questions: How important are commodity options relative to futures in the price formation process? Who is responsible for such price discovery—speculators or hedgers? Using two years of recent intraday data and conventional and new empirical techniques (i.e., the Hasbrouck [1995] information share [IS], the Gonzalo and Granger [1995] component share (CS) and the Yan-Zivot [2010] and Putniņš [2013] information leadership share [ILS]), this study provided a unique and comprehensive examination of price discovery in the agriculture (corn and soybeans), energy (crude oil and natural gas) and metal (gold and silver) commodity futures and options markets. These commodities are selected because they have the largest futures and options trading volume on the CME at the end of 2017. Following the significant growth in trading resulting from electronic trading and improved liquidity, this research focuses on a more recent sample period to report findings that are relevant and have implications for market participants and regulators today. Two approaches, developed by Working (1953) and Lucia and Pardo (2010), were used to identify speculation from hedging activity.

The first key result is that, although futures markets are more liquid in terms of quoted spreads and trading volume, more recently, US commodity options led futures in averages by reflecting new information in the crude oil, gold, silver, corn and soybean markets. This finding highlights the importance of this instrument type, due to recent developments in electronic platforms that help traders initiate a variety of real time options trading strategies across a larger number of assets. Further, no margin requirements are required for long strategies, nor is there the right to trade (rather than obligation). It is likely that these features make options attractive to commodity derivative traders (e.g., hedge funds and HF traders) and such trading activity results in options becoming an important venue for informed trading. It should be noted that, although options lead futures in averages, both markets make a meaningful contribution to price discovery.

In support of these findings, the second key result arises from multivariate analyses. This indicates that a significant determinant for why crude oil, gold, silver, corn and soybean options lead their corresponding futures in price discovery is increased speculation in commodity derivative markets. This finding is robust, as it uses two different measures of speculation. The first incorporates hedging and speculative open interest, which is defined by the US CFTC (Working, 1960). The second is more general and measures incorporating commodity derivatives open interest and trading volume (Lucia & Pardo, 2010).

Prior studies found that hedgers primarily trade in futures markets. Although speculators trade options, futures and combined positions (e.g., protective puts), they prefer options (Benet & Luft, 1995; Lien & Tse, 2000; Tokic, 2013). If there are more hedgers trading in futures markets, this increases the relative proportion of speculators to hedgers in options markets and consistent with the theory of backwardation (which suggests that speculators contribute to price discovery; Chen et al., 2016). As a result, the expectations are larger options price discovery shares because by definition, options *ILS* and futures *ILS* sum to one, therefore, if options *ILS* increases, futures *ILS* decreases. Consistent with this idea, it is found that increases in the relative proportion of speculation in options markets are a key driver for options markets to reflect new information before futures markets.

In addition, it is found that narrower spreads in crude oil and gold options are associated with increases in the options market's contribution to price discovery. This indicates that informed traders trade strategically and choose a relatively more liquid market to utilise their information.

This research contributes to two areas of the existing literature. First, the analysis contributes to the literature on how derivative instruments affect the nature of price discovery. Several studies examine price discovery in commodity futures markets with mixed findings. Garbade and Silber (1983) reported that wheat, corn and orange juice futures lead spot markets. Hauptfleisch, Putniņš and Lucey (2016) reached a similar conclusion in gold futures markets. In contrast, Dimpfl et al. (2017) found that most price discovery occurs in the underlying corn, wheat, soybeans and cattle markets. To the best of our knowledge, Boyd and Locke (2014) provided the only other examination of price discovery in commodity futures and options markets using data during the period from

2005 to 2007. The authors' sampled options prices at 15-minute intervals and found that most price discovery occurs in natural gas futures contracts. This research contributes to the literature by reporting new evidence about price discovery in six commodity futures options markets and for a more recent time period. In addition, price discovery measures are estimated at lower latencies and using new price discovery techniques that account for the noise differential between derivative markets. This approach results in a more accurate measurement of price discovery, allowing for more precise inferences to be made regarding its determinants and time series patterns.

Second, the findings contribute to the literature that examines the role of hedgers and speculators in commodity derivative markets. On the one hand, the dispersion theory suggests that futures speculation is undertaken by uninformed individuals, whereas, hedging activity is conducted by informed parties (e.g., bank, dealers, etc) (Wang, 2004). In contrast, the theory of backwardation suggests that speculative activity contributes to price discovery. Our findings provide support to the theory of backwardation. This analysis adds to this literature by documenting the role of hedgers and speculators in the futures price formation process. This research also uniquely reports new evidence about the role of hedgers and speculators in the futures options price discovery process.

4.2 Commodity Derivative Quoting Activity

To understand price dynamics in commodity derivatives, it is first necessary to understand how prices of commodity futures and options are quoted. Commodity options on the CME are traded through market makers, who receive an RFQ from traders. This is a first distinction to equity options, which follows stock prices according to auto-quoting techniques. According to CME, RFQ allows market participants to remain completely anonymous while the market remains liquid, as market makers are incentivised to quote competitively.[17] Grünbichler, Longstaff and Schwartz (1994) argued that informed traders are drawn to markets that are more anonymous. The authors explained that the shift to screen trading and higher liquidity in all markets leads to decreased transparency, as informed trades camouflage as liquidity traders.

[17] More information about RFQ can be found on https://www.cmegroup.com/education/request-for-quote.html.

Table 4.1 follows Ellis, Michaely and O'Hara (2000) and decomposes the trading activity with respect to the prevailing quotes for futures (Panel A) and options (Panel B). It is found that options markets have higher quote-to-trade ratios (QTT), indicating greater algorithmic activity. Crude oil options have a QTT of over 700, while futures show just below 12. Notably, only futures contracts with the highest volume are used each day, while options are filtered (see Section 3.2 for a description). Further, approximately 80 percent of trades occur at the bid or ask quotes, whereas this figure is close to 100 percent in futures (except natural gas futures). The share of trades outside the prevailing quotes is between four and five percent for agriculture futures options and approximately one percent for the remaining options. Except for corn, futures witness trades outside the quotes in approximately two percent of trades. Therefore, options markets have more mid-quote trades and trades occurring somewhere between the bid and ask prices.

< Table 4.1 here >

4.3 Data

Intraday commodity futures and options quote data are obtained from the *Thomson Reuters Tick History* (*TRTH*) database provided by the Securities Industry Research Centre of Asia–Pacific during a 24-month period between January 2016 and December 2017. Price discovery is examined in the following markets: corn, soybeans, crude oil, natural gas, gold and silver. These commodities were selected because they had the highest daily trading volume at the end of 2017.[18]

Table 4.2 presents contract specifications for each commodity (including futures or options trading symbols, futures unit or contract size and futures or options minimum tick size). Corn and soybean derivatives are traded on the CBOT, crude oil and natural gas derivatives are traded on the New York Mercantile Exchange (NYMEX) and gold and silver derivatives are traded on the Commodity Exchange (COMEX). All futures options are American style; the exercise of a call (put) option results in a long (short) position in

[18] See https://www.cmegroup.com/trading-hours.html for details of trading hours for each market.

the underlying futures contract. Upon the expiration of the futures contract, the short position is required to physically deliver the underlying commodity.

< Table 4.2 here >

Table 4.3 reports the descriptive statistics of the futures and options daily trading activity. Futures volume and open interest are generally larger than the corresponding options figures. Corn and crude oil have the highest average daily futures trading volume (open interest) with over 500,000 (2,000,000) contracts traded in corn and almost 800,000 (1,100,000) contracts traded in crude oil. Silver and soybeans have the lowest futures volumes and open interest. Similar trading activity is observed for commodity options; an average of approximately 180,000 crude oil and 4,000 silver options are traded per day.

< Table 4.3 here >

In comparison, daily futures quoted spreads are considerably smaller than options quoted spreads.[19] Across all commodities, average futures (options) quoted spreads range between 0.005 and 0.548 (0.258 and 10.691). For example, mean options spreads are approximately 20 times larger than futures spreads in the gold, silver and soybean markets.

The combination of lower trading activity and wider quoted spreads in options indicates that options are considerably less liquid than futures. Therefore, options prices are likely to be noisier, affecting the ability of conventional price discovery measures to accurately estimate price discovery (Putniņš, 2013; Yan & Zivot, 2010).[20]

4.4 Futures and Options-implied Price Series

To compute price discovery measures, futures and options-implied price series must be constructed. For each commodity every day, one seamless futures price time series is obtained by rolling across futures contracts with the highest volume. When constructing

[19] Daily quoted spreads (expressed in basis points) are defined as the average of the ratio of the difference between the ask and bid price, divided by the midpoint price.
[20] Boyd and Locke (2014) reach a similar conclusion. They also reported that commodity options prices are noisier than futures prices.

the futures price series, contracts with bid ask spreads greater than 50 percent are omitted. In contrast, the options data comprise several billion observations from contracts with different strike prices and expiration dates. As a result, following Hao (2016), options with zero open interest, a bid ask spread percent of greater than 50 percent and time to maturity of less than five days and more than 90 days are omitted.

Using a similar approach to Muravyev, Pearson and Broussard (2013), options-implied futures price series are calculated. Eq. (4.1) represents the European put-call parity relation to calculate the options-implied futures price for a given put-call pair:

$$F_t e^{-r(T-t)} = C_t(K,T) - P_t(K,T) + PV_t\big(D(t,T)\big) + Ke^{-r(T-t)}, \tag{4.1}$$

where F_t is the futures price at time t, $C_t(K,T)$ and $P_t(K,T)$ are the call and put option prices with strike price K and expiry date T, r is the continuously compounded London Interbank Offered Rate per annum and $T-t$ is the time to maturity.[21] Commodity options are American in style. As a result, Eq. (4.1) is adjusted to capture the ability to exercise early. Denoting the early exercise premium by $v_t(K,T)$, the result is:

$$F_t e^{-r(T-t)} + v_t(K,T) = C_t(K,T) - P_t(K,T) + Ke^{-r(T-t)}, \tag{4.2}$$

where $v_t(K,T)$ is the early exercise premium at time t. The error from the put call parity relation at every quote update is given by:

$$\varepsilon_t = C_t(K,T) - P_t(K,T) + Ke^{-r(T-t)} - F_t e^{-r(T-t)}. \tag{4.3}$$

The early exercise premium is calculated as the average error term (i.e., Eq. [4.3]) for each day. Following this, the options-implied futures bid and ask price at time t are given by:

$$Implied\ Bid_t(K,T) = e^{r(T-t)}[{C_t}^{Bid}(K,T) - {P_t}^{Ask}(K,T) + Ke^{-r(T-t)} - v_t(K,T)], \tag{4.4}$$

$$Implied\ Ask_t(K,T) = e^{r(T-t)}[{C_t}^{Ask}(K,T) - {P_t}^{Bid}(K,T) + Ke^{-r(T-t)} - v_t(K,T)]. \tag{4.5}$$

[21] The LIBOR rate is obtained from the Federal Reserve Bank of St. Louis.

4.5 Price Discovery Measures

Using the futures and options-implied prices, conventional measures of price discovery are estimated. Following Hasbrouck (1995) and Gonzalo and Granger (1995), the *IS* and *CS* measures of price discovery are computed. Both *IS* and *CS* reportedly capture two components: 1) the permanent component (innovations in the fundamental value) and 2) the temporary component (noise). Each day, for each put call pair, using prices sampled in event time-sampling frequency, *IS* and *CS* is estimated from the parameter estimates and then reduced from errors of a VECM with 200 lags:

$$\Delta p_{1,t} = \alpha_1(p_{1,t-1} - p_{2,t-1}) + \sum_{i=1}^{200} \gamma_i \Delta p_{1,t-i} + \sum_{j=1}^{200} \delta_i \Delta p_{2,t-j} + \varepsilon_{1,t}, \quad (4.6)$$

$$\Delta p_{2,t} = \alpha_2(p_{1,t-1} - p_{2,t-1}) + \sum_{k=1}^{200} \varphi_i \Delta p_{1,t-k} + \sum_{m=1}^{200} \phi_i \Delta p_{2,t-m} + \varepsilon_{2,t}, \quad (4.7)$$

where $\Delta p_{i,t}$ is the change in the log mid-quote price of asset i at time t. The normalised orthogonal to the VECM coefficients allows for the computation of *CS*:

$$CS_1 = \gamma_1 = \frac{\alpha_2}{\alpha_2 - \alpha_1}, \ CS_2 = \gamma_2 = \frac{\alpha_1}{\alpha_1 - \alpha_2}. \quad (4.8)$$

Using the covariance matrix of the VECM error terms (i.e., Eq. [4.9]) and its Cholesky factorisation (Eq. [4.10]) allows for the computation of *IS* (Eq. [4.11]):

$$\Omega = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \quad (4.9)$$

$$\Omega = MM', where\ M = \begin{pmatrix} m_{11} & 0 \\ m_{12} & m_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2(1-\rho^2)^{1/2} \end{pmatrix}. \quad (4.10)$$

$$IS_1 = \frac{(\gamma_1 m_{11} + \gamma_2 m_{12})^2}{(\gamma_1 m_{11} + \gamma_2 m_{12})^2 + (\gamma_2 m_{22})^2}, IS_2 = \frac{(\gamma_2 m_{22})^2}{(\gamma_1 m_{11} + \gamma_2 m_{12})^2 + (\gamma_2 m_{22})^2}. \quad (4.11)$$

Following Baillie et al. (2002), the ordering of prices in the VECM is accounted for by estimating *IS* under both orderings and taking the simple average.

Prior studies use *IS* and *CS* to measure what market is the first to reflect new information (i.e., the permanent component of price changes). However, Yan and Zivot (2010), Putniņš (2013) showed that *IS* and *CS* capture both permanent and temporary (i.e., noise) components of price discovery. Of particular concern, both *IS* and *CS* overstate the contribution of price discovery of the market with lower levels of noise. This is an issue for prior studies examining price discovery between underlying and derivative markets, which influences price discovery estimated between futures and options markets. Table

4.2 indicates that commodity futures are more liquid, or less noisy, than options markets. Consequently, the *ILS* developed by Yan and Zivot (2010) and Putniņš (2013) provides a solution to this problem. *ILS* combines *IS* and *CS* to cancel out the dependence on noise, resulting in a measure that captures permanent price changes (information leadership) only:

$$ILS_1 = \frac{\left|\frac{IS_1 CS_2}{IS_2 CS_1}\right|}{\left|\frac{IS_1 CS_2}{IS_2 CS_1}\right| + \left|\frac{IS_2 CS_1}{IS_1 CS_2}\right|}, ILS_2 = \frac{\left|\frac{IS_2 CS_1}{IS_1 CS_2}\right|}{\left|\frac{IS_1 CS_2}{IS_2 CS_1}\right| + \left|\frac{IS_2 CS_1}{IS_1 CS_2}\right|}. \quad (4.12)$$

All three price discovery measures are computed by obtaining estimates that range between [0,1]. Values above 0.5 indicate that one market leads price discovery.

4.6 Identification of Speculators and Hedgers

Following Peck (1980), Sanders, Irwin and Merrin (2010) and Dimpfl, Flad and Jung (2017), speculators and hedgers are identified in commodity options and futures markets using the *WorkingT* index (Working, 1953). The intuition behind *WorkingT* is that it captures the excess speculative trading activity required to balance the difference in unmatched long and short hedging activity. To estimate *WorkingT*, weekly data contained within the Commitment of Traders reports released by the CFTC is used. This identifies the closing total outstanding contracts (i.e., open interest) of commercial and non-commercial options and futures traders when the literature classes commercial traders as hedgers and non-commercial traders as speculators (e.g., Sanders et al., 2009; Tokic, 2013).[22]

Table 4.4 reports the average weekly net open interest (i.e., long minus short) combined across futures and options for hedgers, speculators and non-reportable parties. For all commodities, net hedgers have a net short position, especially when trading derivatives on crude oil and corn in which traders are likely to protect the sale of these commodities. In contrast, with the exception of corn, net speculators have net long positions in futures

[22] COT data are obtained from http://www.cftc.gov. The CFTC classifies traders as commercial or non-commercial on a yearly basis. Non-reportable open interest, which identifies small traders, is omitted. The CFTC only reports the combined futures and options open interest data.

and options, especially in crude oil derivatives. These findings indicate a bullish sentiment among traders in most commodities examined in this study.

< Table 4.4 here >

Using Eq. (4.13), *WorkingT* is calculated for each commodity each week:

$$WorkingT = \begin{cases} 1 + \frac{S\_Short}{(H\_Long + H\_Short)} \; where \; H\_Short \; \geq H\_Long \\ 1 + \frac{S\_Long}{(H\_Long + H\_Short)} \; where \; H\_Short < H\_Long \end{cases}, \qquad (13)$$

where *S_Long* (*S_Short*) is the long (short) options and futures open interest of non-commercial traders (speculators) and *H_Long* (*H_Short*) is the long (short) options and futures open interest of commercial traders (hedgers). *WorkingT* values greater than one indicate excessive speculative trading activity, which is required to cover hedging positions. If *WorkingT* is equal to one, this suggests there is no excess speculative trading activity.

Table 4.5 reports descriptive statistics for the *WorkingT* measure of excess speculation. Evidence of excessive speculation is observed in all six commodity markets—*WorkingT* mean values ranging from 1.090 to 1.567. In all cases, similar mean and median values are reported. The largest (smallest) mean levels of variation in excess speculation occur in natural gas (soybean) markets. For example, a mean *WorkingT* value of 1.567 indicates that 56.7 percent of speculative positions exceed the amount required to cover hedging activity in natural gas markets. The subsequent analysis examines whether such excessive speculation in commodities is associated with price discovery in options or futures markets.

< Table 4.5 here >

4.7 Price Discovery Estimates

For each commodity, price discovery is estimated using conventional and new empirical techniques. Table 4.6 Panel A reports average daily values estimated using prices sampled in event time and using 200 lags in the VECM. In all cases, both futures *IS* and *CS* (values slightly larger than 0.5 and approximately ranging from 0.5 to 0.6) indicate that commodity futures are, on average, the first to reflect new information relative to options markets. Figure 4.1 reports the time series of options price discovery between 2016 and 2017. Downtrends were observed in both the options *IS* and *CS* during the second half of the sample period for crude oil, natural gas, gold and silver (noting that futures *IS* one minus options *IS*, which is similar for futures *CS*).

< Table 4.6 here >

In contrast, *ILS* suggests that, on average, commodity options lead futures in impounding new information, with average values ranging from 0.576 to 0.623. The exception is natural gas, in which options *ILS* is 0.477. Using futures *IS*, Boyd and Locke (2014) also found that natural gas futures lead options in price discovery between 2005 and 2007.

Figure 4.1 shows uptrends in options *ILS* in the second half of the sample period for energy and metal commodities. Such trends are potentially consistent with an increased preference for informed trading in options, due to developments in electronic and real time trading, the ability to implement various options trading strategies, weekly options contracts and the CME's closure of the futures trading pit (Sammann, 2015; Simon, 2014). Options and futures *ILS* are relatively constant for agriculture commodities. Given that *ILS* is insensitive to noise differentials between the two markets, it is likely to provide the most accurate portrayal of price discovery between the two markets. These findings indicate that both options and futures are important venues for informed trading, with options becoming the dominant venue during more recent time periods.

< Figure 4.1 here >

4.8 Price Discovery Drivers in Futures and Options Markets

This section examines whether the trading activity of speculators or hedgers contributes to price discovery in futures and options markets. In a similar vein to Chakravarty et al. (2004) and Boyd and Locke (2014), the determinants of options price discovery are estimated using the following time-series regression with clustered standard errors by month:

$$ILS_{i,t} = \beta_0 + \beta_1 METRIC_{i,t} + \beta_2 \frac{OptVol_{i,t}}{FutVol_{i,t}} + \beta_3 \frac{OptOI_{i,t}}{FutOI_{i,t}} + \beta_4 \frac{OptQSpr_{i,t}}{FutQSpr_{i,t}} + \beta_5 Skew_{i,t} + \beta_6 Vola_{i,t} + \varepsilon_{i,t}, \tag{4.14}$$

where *ILS* is the Yan and Zivot (2010) and Putniņš (2013) *ILS* in week *t* with prices sampled in event time and 200 lags in the VECM. Due to the sensitivity of noise inherent in conventional measures of price discovery, *ILS* is used to measure price discovery. To obtain a weekly *ILS* value, the daily value is taken and averaged across the week (there are approximately 104 weeks during the two-year sample period).[23] In Model 1, *METRIC* is *WorkingT* is used, which captures speculative activity in options and futures markets. The independent variables include ratios of options and futures trading volume (*OptVol*/*FutVol*), options and futures open interest (*OptOI*/*FutOI*) and options and futures quoted bid-ask spreads (*OptQSpr*/*FutQSpr*). Skewness of futures returns (*Skew*) follow Fernandez-Perez et al. (2018) and 20-day intraday return volatility (*Vola*) are also used.[24] A positive (negative) coefficient on the *METRIC* variable indicates that increases (decreases) in excess speculation are associated with increases (decreases) in the options market's share of price discovery.

For each commodity, Table 4.7 Model 1 reports the regression findings. Panel A (B) reports results in which *ILS* is measured using prices sampled in event time (at one second) frequency. Across Panels A and B, for all commodities, a positive and significant relationship is reported between relative options *ILS* and *WorkingT*. The exceptions are silver in Panel A and crude oil in Panel B. Increases in excess speculation in the soybeans and gold markets result in the largest increases in relative options *ILS* (e.g., Panel A *METRIC* coefficient estimates of 0.315 and 0.269). These results are consistent with the theory of backwardation; that is, increases in excess speculation in options and futures energy, metal and agriculture markets are associated with an increase in the share of price discovery occurring in options markets. Such findings are consistent with the recent growth in futures options trading activity resulting from the development of electronic platforms, allowing for real time trading, improvements in liquidity and the ability and

[23] For each commodity, using an Augmented Dickey Fuller test, the hypothesis that options ILS is non-stationary (p-values less than 0.01) is rejected.

[24] Weekly *OptVol* (*FutVol*) is calculated as the sum of options (futures) volume across all contracts per week (weekly *OptOI* and *FutOI* is calculated in a similar vein). Weekly *OptQSpr* (*FutQSpr*) is calculated as the daily *OptQSpr* (*FutQSpr*) values averaged across the week (weekly *Skew* and *Vola* are calculated in a similar vein).

flexibility to trade complex and various options strategies (Sammann, 2015; Simon, 2014; Wolf, 1982).

< Table 4.7 here >

In unreported findings, if *CS* is used as the dependent variable in Eq. (4.14), it is found that increased speculation in commodity derivative markets (as captured by the *WorkingT* variable) is associated with *lower* options *CS* in the gold, silver and soybean markets. Similar results are found using *IS* as the dependent variable. Given that *IS* and *CS* are partly influenced by noise, such findings indicate that there is an increase in the relative noise in options market quotes when more speculative activity takes place in commodity derivative markets. This results in lower *IS* and *CS* values, as they are biased downwards by noise.

Using relative options and futures bid-ask spreads, this research examines whether informed traders prefer to trade strategically in the more liquid market and maximise the value of their information (Chakravarty et al., 2004; Fleming et al., 1996). Consistent with this idea, it was found that relatively narrower spreads in crude oil and gold options are associated with increases in the options market's contribution to price discovery. Conversely, a significant and positive relationship was observed between relative options price discovery and relative options and futures bid-ask spreads in the natural gas and corn markets. This relationship is consistent with more informed trading and speculation in options markets, which increases adverse selection risks in options markets, increasing relative options spreads (Glosten & Milgrom, 1985).

4.9 Robustness

In addition to the robustness tests described in the preceding sections of the paper, additional tests were conducted that use alternative ways of measuring speculation and hedging activity in commodity derivative markets.

First, an alternative approach was used to capture speculation and hedging in commodity derivative markets. Methods by Lucia and Pardo (2010) were used to infer daily hedging and speculative activity using options and futures open interest and volume. Each day, the following was estimated:

$$R_t = \frac{\Delta OI_t}{V_t}\ x\ 1{,}000, \tag{4.15}$$

where ΔOI_t is the daily change in options and futures open interest on day t and V_t is the total options and futures volume. The assumptions behind this measure is that new options and futures positions are a proxy for hedging activity; actual traded volume is a proxy for speculative activity. Higher (lower) values of R capture hedging (speculative) activity.

Table 4.5 reports descriptive statistics for the R measure. Mean and median values are similar for each commodity, with means ranging between 1.249 and 3.747. Based on the change in daily open interest, the largest levels of hedging are observed in metals and the lowest in soybean markets.

Qualitatively similar findings are reported in Table 4.7 Panels A and B Model 2, if Eq. (4.14) is re-estimated using daily observations (instead of weekly observations) and when *METRIC* is the R measure of speculative or hedging activity (instead of *WorkingT* in Model 1). It can be observed that increases in speculation in options and futures markets (as captured by a negative coefficient estimate on the R variable) is associated with increased options price discovery (or decreased price discovery in commodity futures) in crude oil, gold, silver, corn and soybean markets. It was observed that the magnitude of speculation in the crude oil and soybean markets results in the largest increases in options price discovery (e.g., Panel A *METRIC* coefficient estimates of –10.310 and –0.541).

A positive relationship is noted between natural gas options *ILS* and the R variable. Ritz (2014) described natural gas prices as unique, as there is no single global market, but regional hubs with significant differences in prices. Prices differ due to varying transport costs, constraints on shipping and irrational producer-selling prices. Differences in market structure, varying prices and an inability for traders to engage in arbitrage trading provide potential explanations for why different and mixed findings were observed when examining speculation and price discovery in natural gas commodity derivatives compared to other commodity markets. Future research could provide further analysis of this.

In unreported findings, it was found that increased speculation captured by *R* in commodity derivative markets is associated with an increase in the relative noise in options quotes (as measured by options *IS* or *CS*).

To examine whether the relative proportion of speculation in options affects the speed of options markets when reflecting new information, Eq. (4.14) was re-estimated when *METRIC* (*WorkingT* or *R*) was defined as the relative proportion of speculation in options markets to the total speculation across options and futures markets.[12] For example, *METRIC* is *WorkingT* calculated using options markets data divided by *WorkingT* calculated using options and futures markets data. Unreported findings show that relative increases in options market speculation (captured by *WorkingT* and *R*) are associated with increased price discovery in all six commodities options markets (the findings in silver markets are the weakest).

4.10 Summary

The motivation for this study arises from developments in the ability to trade commodity derivatives and resulting increases in trading activity, a lack of understanding regarding the role of futures options in price discovery and the continued debate regarding the role of hedgers and speculators in the price formation of commodity derivatives. As a result, this research provides a unique examination of price discovery in six different commodity futures options markets. The role of speculators and hedgers in the price discovery process is analysed. Further, the trading activity is disentangled with respect to the prevailing quotes in commodity futures and options.

Using intraday data, conventional and new empirical measures of price discovery are estimated, noting that they measure different components of price discovery. *IS* and *CS* are measures of the relative level of noise between the two markets; *ILS* captures the relative speed with which each market reflects new information. This is the traditional focus of using such empirical measures of price discovery. Further, using open interest and volume data, two measures are used—*WorkingT* and *R*—to identify speculation and hedging activity in commodity derivative markets.

Despite lower levels of liquidity, it was found that, on average, options lead futures in reflecting new information in the crude oil, gold, silver, corn and soybean markets. Such findings are consistent with recent developments in trading platforms, allowing various market participants the flexibility and ease to trade a variety of options strategies. Although most price discovery occurs in options (approximately 0.55–0.6), a large fraction of price discovery occurs in futures markets (approximately 0.4–0.45).

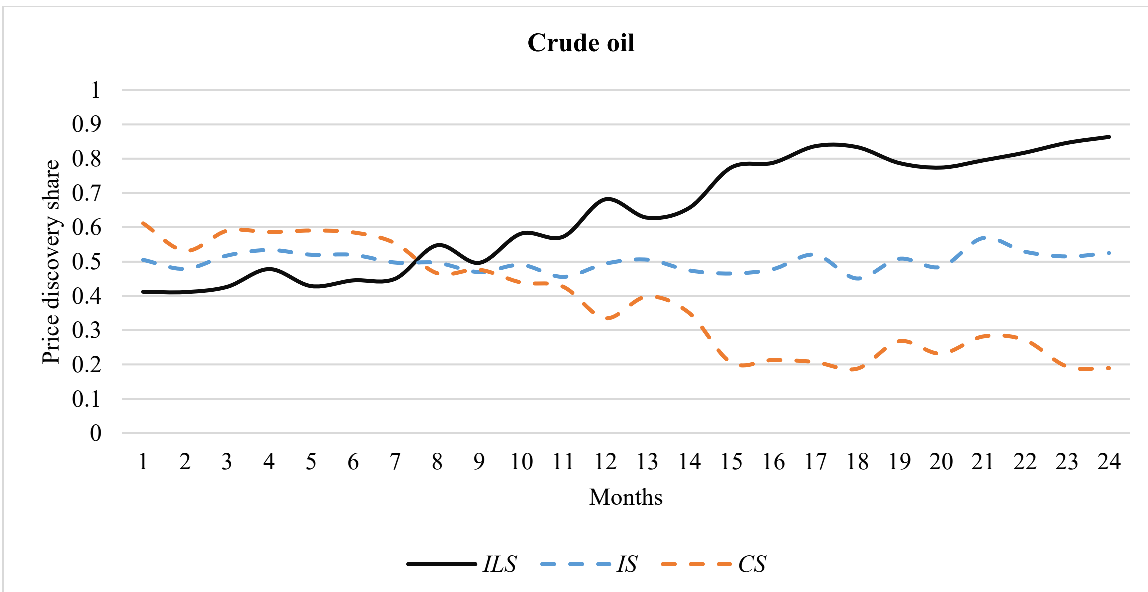
In addition, it was found that speculation is a significant determinant of price discovery in commodity derivatives. Specifically, it was reported that increased speculation in commodity derivatives is associated with increases in the options market's contribution to price discovery. This result occurs in crude oil, gold, silver and soybean markets. By examining price discovery between markets of differing levels of liquidity, the findings of this study highlight the importance of accounting for noise differential using empirical measures such as *ILS*.

Future research can examine the level of price discovery in other commodity and derivative markets. Further, with the increasing popularity of futures options, increases in electronic trading, changing regulation and market structure, further research can examine the drivers of trading activity and price discovery, especially as other countries and exchanges introduce commodity derivative trading platforms.

Figure 4.1: Price discovery over time

This figure reports the average options market price discovery shares for each commodity between 1 January 2016 and 31 December 2017. Price discovery shares are calculated using the Hasbrouck (1995) *IS*, Gonzalo-Granger (1995) *CS* and Yan-Zivot (2010) and Putniņš (2013) *ILS*. Price discovery measures are estimated using prices sampled in event time and using 200 lags in the VECM.

Energy



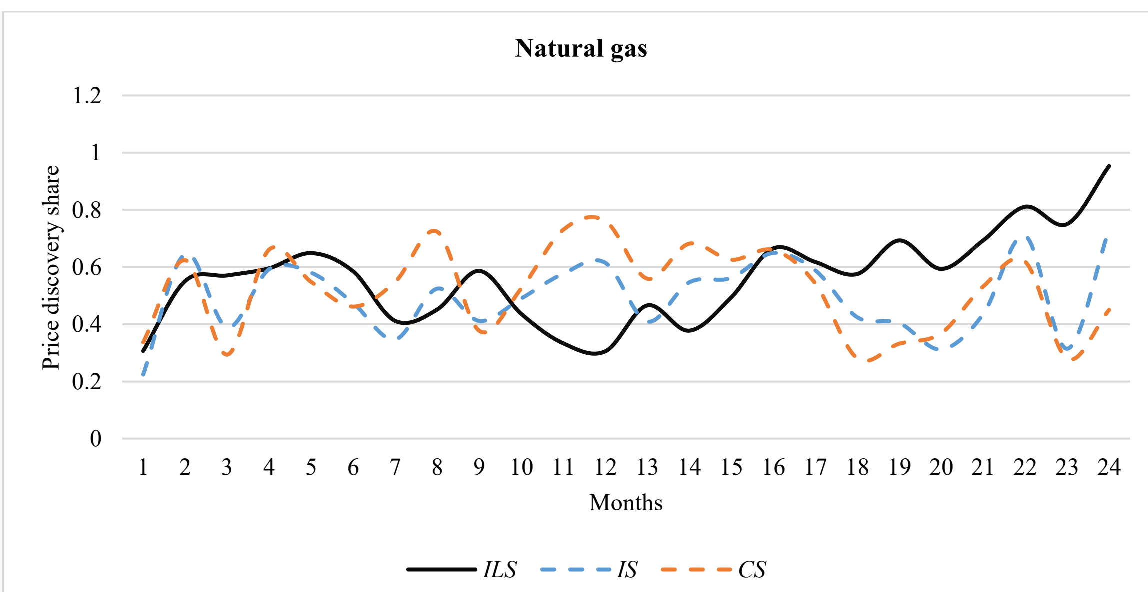
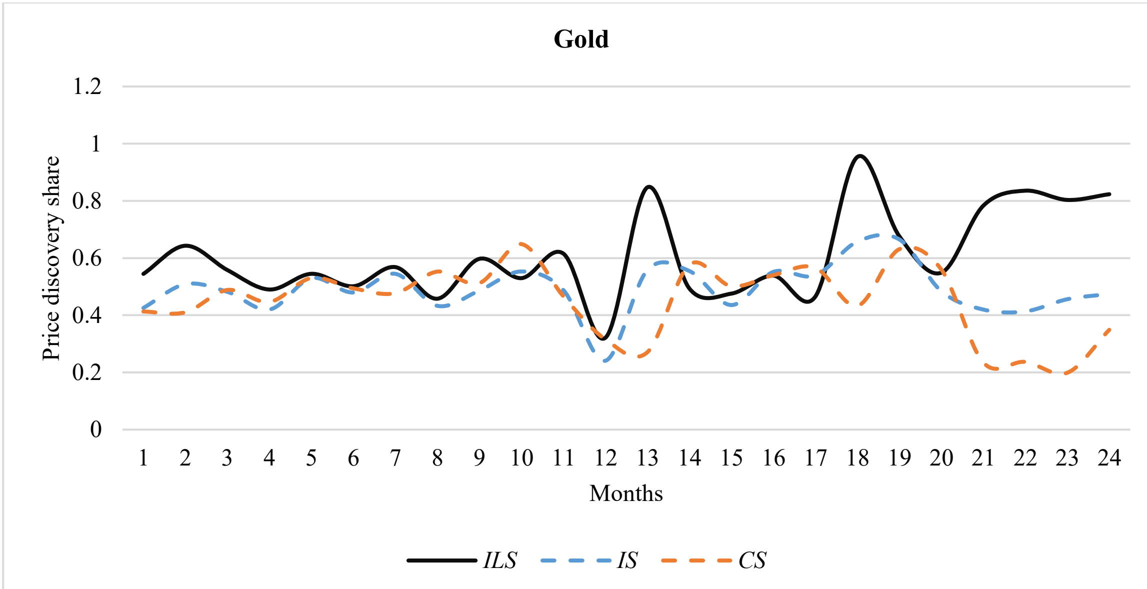


Figure 4.1: Price discovery over time (continued)

Metals



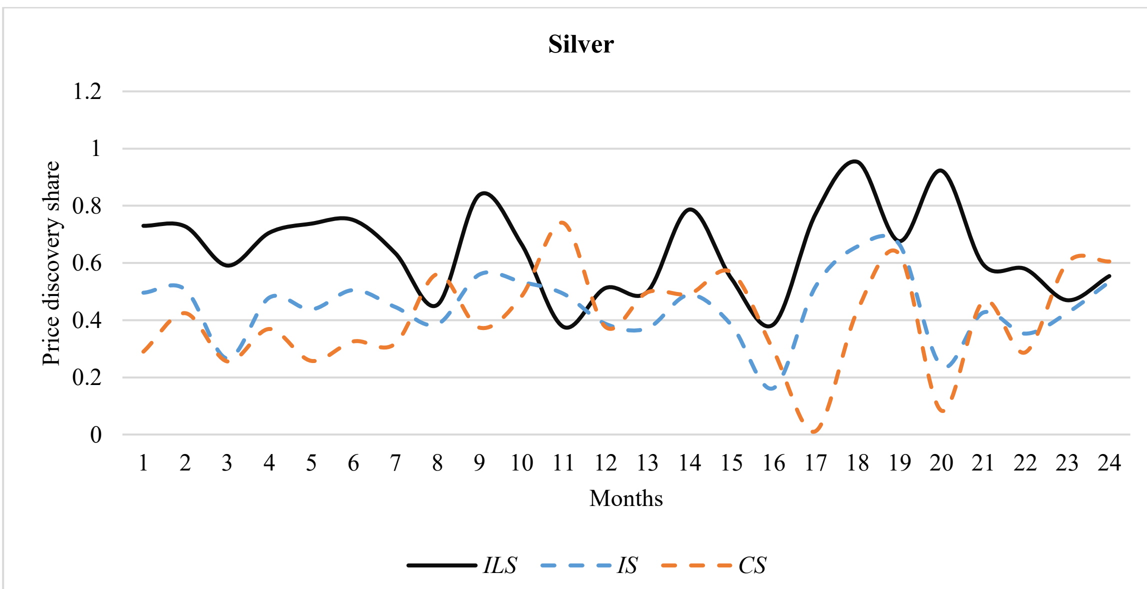
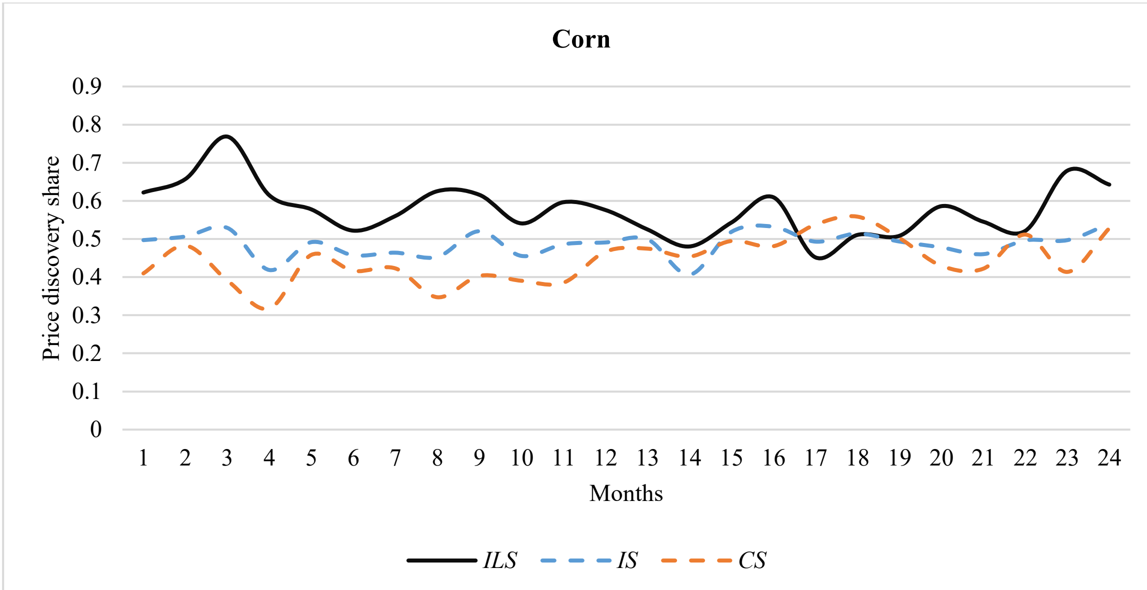


Figure 4.1: Price discovery over time (continued)

Agriculture



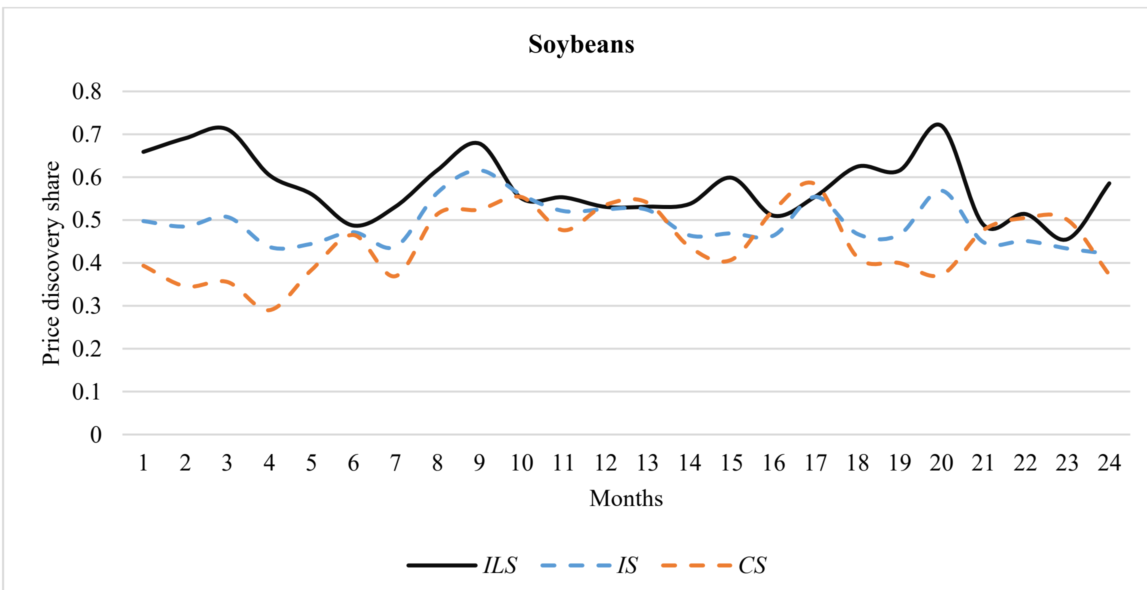


Table 4.1: Distribution of trades with respect to quotes

This table reports the number of trades with respect to the prevailing quotes. Panel A shows data for futures and Panel B shows data for options. The QTT is calculated as the number of quotes divided by the number of trades.

Panel A: Futures						
Contract	**Crude oil - futures**	**Natural gas - futures**	**Gold - futures**	**Silver - futures**	**Corn - futures**	**Soybeans - futures**
Total trades	**81,553,374**	**16,824,663**	**39,488,664**	**13,080,204**	**7,370,076**	**10,809,189**
Total quotes	**972,319,037**	**164,667,954**	**453,354,129**	**212,032,800**	**115,171,244**	**156,220,432**
Quote-to-trade ratio	*11.92*	*9.79*	*11.48*	*16.21*	*15.63*	*14.45*
Trades inside quotes	6,195,717	3,321,651	245,866	167,771	200,362	200,154
% of all trades	*7.60%*	*19.74%*	*0.62%*	*1.28%*	*2.72%*	*1.85%*
Trades bid	37,064,112	6,567,924	19,014,076	6,352,476	3,709,439	5,247,565
% of all trades	*45.45%*	*39.04%*	*48.15%*	*48.57%*	*50.33%*	*48.55%*
Trades at ask	37,177,048	6,599,785	19,074,105	6,400,424	3,424,875	5,140,041
% of all trades	*45.59%*	*39.23%*	*48.30%*	*48.93%*	*46.47%*	*47.55%*
Trades at mid-quote	30,944	19,049	6220	6775	2842	3674
% of all trades	*0.04%*	*0.11%*	*0.02%*	*0.05%*	*0.04%*	*0.03%*
Trades outside quotes	1,116,497	335,303	1,154,617	159,533	35,400	221,429
% of all trades	*1.37%*	*1.99%*	*2.92%*	*1.22%*	*0.48%*	*2.05%*
thereof above ask	555,565	163,173	566,414	75,852	16,965	108,609
% of all trades	*0.68%*	*0.97%*	*1.43%*	*0.58%*	*0.23%*	*1.00%*
thereof below bid	560,932	172,130	588,203	83,681	18,435	112,820
% of all trades	*0.69%*	*1.02%*	*1.49%*	*0.64%*	*0.25%*	*1.04%*

Table 4.1: Distribution of trades with respect to quotes (continued)

Panel B: Options												
Contract	**Crude oil - calls**	**Crude oil - puts**	**Natural gas - calls**	**Natural gas - puts**	**Gold - calls**	**Gold - puts**	**Silver - calls**	**Silver - puts**	**Corn - calls**	**Corn - puts**	**Soybean - calls**	**Soybean - puts**
Total trades	**557,788**	**567,227**	**271,245**	**240,758**	**663,900**	**506,776**	**227,438**	**149,412**	**665,884**	**476,574**	**886,272**	**781,942**
Total quotes	**391,260,490**	**399,149,059**	**63,935,751**	**53,426,507**	**157,158,080**	**128,052,884**	**39,126,048**	**39,914,905**	**29,610,950**	**38,207,205**	**175,053,632**	**146,790,681**
Quote-to-trade ratio	701.45	703.68	235.71	221.91	236.72	252.68	172.03	267.15	44.47	80.17	197.52	187.73
Trades inside quotes	139,343	136,033	54,817	52,346	170,003	135,054	47,327	29,127	106,896	85,602	159,187	113,979
% of all trades	*24.98%*	*23.98%*	*20.21%*	*21.74%*	*25.61%*	*26.65%*	*20.81%*	*19.49%*	*16.05%*	*17.96%*	*17.96%*	*14.58%*
Trades bid	203,012	208,899	102,520	88,997	234,148	177,470	86,041	57,105	258,126	182,053	339,353	312,125
% of all trades	*36.40%*	*36.83%*	*37.80%*	*36.97%*	*35.27%*	*35.02%*	*37.83%*	*38.22%*	*38.76%*	*38.20%*	*38.29%*	*39.92%*
Trades at ask	210,478	217,869	112,200	98,267	250,473	187,234	91,819	61,535	264,503	185,267	345,917	322,780
% of all trades	*37.73%*	*38.41%*	*41.36%*	*40.82%*	*37.73%*	*36.95%*	*40.37%*	*41.18%*	*39.72%*	*38.87%*	*39.03%*	*41.28%*
Trades at mid-quote	5,470	6,199	2,425	2,296	3,599	2,602	570	339	38,491	18,583	30,628	27,273
% of all trades	*0.98%*	*1.09%*	*0.89%*	*0.95%*	*0.54%*	*0.51%*	*0.25%*	*0.23%*	*5.78%*	*3.90%*	*3.46%*	*3.49%*
Trades outside quotes	4,955	4,426	1,708	1,148	9,276	7,018	2,251	1,645	36,359	23,652	41,815	33,058
% of all trades	*0.89%*	*0.78%*	*0.63%*	*0.48%*	*1.40%*	*1.38%*	*0.99%*	*1.10%*	*5.46%*	*4.96%*	*4.72%*	*4.23%*
Thereof above ask	2,506	2,180	1,070	681	5,148	3,270	1,367	775	25,719	11,571	27,804	18,645
% of all trades	*0.45%*	*0.38%*	*0.39%*	*0.28%*	*0.78%*	*0.65%*	*0.60%*	*0.52%*	*3.86%*	*2.43%*	*3.14%*	*2.38%*
Thereof below bid	2,449	2,246	638	467	4,128	3,748	884	870	10,640	12,081	14,011	14,413
% of all trades	*0.44%*	*0.40%*	*0.24%*	*0.19%*	*0.62%*	*0.74%*	*0.39%*	*0.58%*	*1.60%*	*2.53%*	*1.58%*	*1.84%*

Table 4.2: Commodity derivative contract specifications

This table reports futures and options contract specifications obtained from the CME for each commodity. The specifications include exchange listing, futures and options contract symbols, futures contract size or unit, futures and options minimum tick size and options style. MMBtu is one million British thermal units.

Contract specifications	Exchange	Futures symbol	Futures unit	Futures minimum tick size	Options symbol	Options minimum tick size	Options style
Energy							
Crude oil (light sweet)	NYMEX	CL	1,000 barrels	1ct/barrel	LO	1ct/barrel	American
Natural gas (Henry hub)	NYMEX	NG	10,000 mmBtu	0.1ct per MMBtu	LNE	0.1ct per MMBtu	American
Metals							
Gold	COMEX	GC	100 troy ounces	10ct per troy ounce	OG	10ct per troy ounce	American
Silver	COMEX	SI	5,000 troy ounces	0.5ct per troy ounce	SO	10ct per troy ounce	American
Agriculture							
Corn	CBOT	ZC	5,000 bushels	0.25ct/bushel	OZC	0.125ct/bushel	American
Soybean	CBOT	ZS	5,000 bushels	0.25ct/bushel	OZS	0.005ct/bushel	American

Table 4.3: Descriptive statistics of futures and options trading activity

This table reports descriptive statistics of futures and options trading activity for each commodity between 1 January 2016 and 31 December 2017. Variables of interest include daily futures returns (*Return*), daily futures return skewness (*Skew*), daily options and futures trading volume in thousands (*OptVol* and *FutVol*), daily options and futures open interest in thousands (*OptOI* and *FutOI*) and daily options and futures bid-ask spread (*OptQSpr* and *FutQSpr*).

	Mean	Median	Std dev.	Mean	Median	Std dev.
		Crude oil			**Natural gas**	
Return	0.149	0.00	0.697	0.905	3.004	0.846
Skew	–0.061	–0.124	0.523	–0.752	–0.465	0.888
FutVol	781.64	138.47	1,686.40	300.87	101.86	467.41
FutOI	1,154.05	590.76	1,374.50	750.09	476.65	771.61
FutQSpr	0.014	0.014	0.024	0.005	0.002	0.022
OptVol	178.66	160.36	83.77	12.12	10.42	7.04
OptOI	2,575.08	2,608.77	401.60	173.27	186.48	64.72
OptQSpr	0.443	0.371	0.331	0.315	0.076	2.171
		Gold			**Silver**	
Return	1.618	0.374	0.465	1.485	2.595	0.713
Skew	–0.785	–0.474	1.200	–0.715	–0.750	1.213
FutVol	340.88	10.16	825.54	141.91	5.87	279.32
FutOI	564.50	92.96	1076.99	256.22	26.03	458.54
FutQSpr	0.096	0.085	0.982	0.013	0.006	0.036
OptVol	29.27	25.11	19.52	3.81	3.32	2.73
OptOI	862.60	935.60	220.33	99.54	104.29	26.44
OptQSpr	1.554	1.144	3.908	0.258	0.223	0.369
		Corn			**Soybean**	
Return	–0.761	2.737	0.529	0.475	1.057	0.537
Skew	–0.587	–0.482	1.102	–0.606	–0.337	0.974
FutVol	534.69	146.91	758.34	254.59	44.14	445.88
FutOI	2,034.04	1,074.87	2321.25	746.55	231.84	1,087.95
FutQSpr	0.050	0.148	0.972	0.548	0.468	0.746
OptVol	74.46	60.04	47.64	72.45	60.25	42.16
OptOI	1,072.83	1,115.14	243.69	754.89	748.98	271.73
OptQSpr	3.350	3.463	2.551	10.691	10.606	6.069

Table 4.4: Net hedging and speculation in commodity derivative markets

This table reports the average weekly net open interest combined across futures and options contracts for each commodity divided into hedging, speculation and non-reportable activity expressed in thousands between 1 January 2016 and 31 December 2017. The CFTC defines net hedging as the difference between producers and merchants long and short positions (i.e., commercial traders), net speculation as the difference between managed money long and short positions (i.e., non-commercial traders) and net non-reportable as the difference between non-reportable long and short positions.

Combined open interest	Net hedging	Net speculation	Net non-reportable
Crude oil	–217.362	233.057	7.122
Natural gas	–46.726	35.309	32.177
Gold	–131.365	147.821	18.446
Silver	–51.450	52.600	13.318
Corn	–235.178	–77.511	–23.178
Soybeans	–122.198	56.966	–49.658

Table 4.5: Speculation and hedging measures

This table reports descriptive statistics of the speculation and hedging measures for each commodity. These measures include *WorkingT*, which captures excess speculation required to meet hedging activity (speculation and hedging activity is defined by the CFTC) and *R*, which captures the percentage of hedging activity (change in open interest) as a fraction of speculative activity (volume) multiplied by 1,000. Weekly (daily) descriptive statistics for *WorkingT* (*R*) are reported. The sample period is between 1 January 2016 and 31 December 2017.

	Mean	Median	Std dev.	Mean	Median	Std dev.
		Crude oil			Natural gas	
WorkingT	1.116	1.100	0.065	1.567	1.591	0.239
R	1.643	1.573	0.606	1.685	1.507	0.947
		Gold			**Silver**	
WorkingT	1.230	1.163	0.167	1.273	1.199	0.207
R	3.747	3.302	2.162	3.630	2.968	2.857
		Corn			**Soybean**	
WorkingT	1.259	1.271	0.090	1.090	1.083	0.057
R	1.905	1.645	1.001	1.249	1.172	0.589

Table 4.6: Price discovery in commodity derivative markets

This table reports the mean options and futures market price discovery shares for each commodity. The price discovery shares are calculated using the Hasbrouck (1995) *IS*, Gonzalo-Granger (1995) *CS* and Yan-Zivot (2010) and Putniņš (2013) *ILS*. Panel A (B) reports price discovery measures in which samples are priced in event time (one second) sampling frequency and 200 lags in the VECM are used. Panel C (D) reports price discovery measures in which samples are priced in event time (one second) sampling frequency and 60 lags in the VECM are used. Grey shading indicates price discovery estimates greater than 0.5. This sample period is between 1 January 2016 and 31 December 2017.

	IS (futures)	*IS* (options)	*CS* (futures)	*CS* (options)	*ILS* (futures)	*ILS* (options)
Panel A: Event time-sampling frequency						
Crude oil	0.502	0.498	0.606	0.394	0.377	0.623
Natural gas	0.518	0.482	0.725	0.275	0.523	0.477
Gold	0.522	0.478	0.570	0.430	0.401	0.599
Silver	0.554	0.446	0.578	0.422	0.387	0.613
Corn	0.521	0.479	0.571	0.429	0.424	0.576
Soybean	0.508	0.492	0.565	0.435	0.421	0.579
Panel B: One-second sampling frequency						
Crude oil	0.520	0.480	0.601	0.399	0.419	0.581
Natural gas	0.553	0.447	0.685	0.315	0.574	0.426
Gold	0.510	0.490	0.566	0.434	0.431	0.569
Silver	0.554	0.446	0.551	0.449	0.416	0.584
Corn	0.501	0.499	0.514	0.486	0.470	0.530
Soybean	0.568	0.432	0.606	0.394	0.468	0.532
Panel C: Event time-sampling frequency						
Crude oil	0.520	0.480	0.595	0.405	0.376	0.624
Natural gas	0.508	0.492	0.593	0.407	0.519	0.481
Gold	0.512	0.488	0.557	0.443	0.410	0.59
Silver	0.548	0.452	0.596	0.404	0.382	0.618
Corn	0.512	0.488	0.568	0.432	0.422	0.578
Soybean	0.516	0.484	0.562	0.438	0.439	0.561
Panel D: One-second sampling frequency						
Crude oil	0.503	0.497	0.594	0.406	0.407	0.593
Natural gas	0.511	0.489	0.647	0.353	0.595	0.405
Gold	0.511	0.489	0.568	0.432	0.421	0.579
Silver	0.533	0.467	0.586	0.414	0.422	0.578
Corn	0.506	0.494	0.565	0.435	0.428	0.572
Soybean	0.516	0.484	0.578	0.422	0.438	0.562

Table 4.7: Speculation in commodity derivative markets

This table reports the coefficient estimates of the determinants of price discovery for each commodity from the following regression:

$$ILS_{i,t} = \beta_0 + \beta_1 METRIC_{i,t} + \beta_2 \frac{OptVol_{i,t}}{FutVol_{i,t}} + \beta_3 \frac{OptOI_{i,t}}{FutOI_{i,t}} + \beta_4 \frac{OptQSpr_{i,t}}{FutQSpr_{i,t}} + \beta_5 Skew_{i,t} + \beta_6 Vola_{i,t} + \varepsilon_{i,t},$$

where *ILS* is options price discovery calculated using the Yan and Zivot (2010) and Putniņš (2013) *ILS* at time *t*. *METRIC* represents each of the following speculation measures: *WorkingT*, which is calculated on a weekly basis, or *R*, which is calculated on a daily basis. The independent variables include ratios of options and futures trading volumes (*OptVol*/*FutVol*), options and futures open interest (*OptOI*/*FutOI*), options and futures bid-ask spreads (*OptQSpr*/*FutQSpr*), skewness of futures returns (*Skew*) and 20-day intraday return volatility (*Vola*). *N* is the number of observations (using weekly observations in Model 1 and daily observations in Model 2). Panel A (B) reports *ILS* in which prices are sampled in event time (one second) sampling frequency and 200 lags in the VECM are used. Standard errors are clustered by month. *t*-statistic values are shown in brackets. ***, ** and * indicate statistical significance at the 1 percent, 5 percent and 10 percent level.

Panel A	Crude oil		Natural gas		Gold		Silver		Corn		Soybeans	
	Model 1: *WorkingT*	Model 2: *R*	Model 1: *WorkingT*	Model 2: *R*	Model 1: *WorkingT*	Model 2: *R*	Model 1: *WorkingT*	Model 2: *R*	Model 1: *WorkingT*	Model 2: *R*	Model 1: *WorkingT*	Model 2: *R*
Intercept	0.341***	2.711***	0.034	0.935	0.490***	2.279***	–0.009	–0.699*	0.024	1.033***	–0.328***	0.548***
	(27.33)	(49.96)	(0.28)	(1.00)	(8.72)	(25.25)	(–0.09)	(–1.66)	(0.88)	(10.32)	(–8.15)	(8.73)
Metric	0.145***	–10.310***	0.181***	6.542***	0.269***	–0.375***	0.006	–0.378**	0.047***	–0.189*	0.315***	–0.541***
	(11.88)	(–23.66)	(2.60)	(3.58)	(5.00)	(–3.66)	(0.09)	(–2.31)	(2.64)	(–1.68)	(8.88)	(–6.23)
OptVol/ FutVol	–2.077***	–0.247***	–3.930	0.681	–0.541	–0.077*	10.246***	10.794***	–1.304***	–1.385***	0.609***	75.236***
	(–10.39)	(–12.39)	(–0.36)	(0.62)	(–1.36)	(–1.93)	(2.72)	(2.87)	(–7.49)	(–8.02)	(5.49)	(6.88)
OptOI/ FutOI	–0.027***	–0.624***	–0.938***	–8.574***	–0.753***	–0.776***	–0.087	0.159	0.283***	0.191*	0.297***	24.071***
	(–4.61)	(–10.69)	(–3.43)	(–3.17)	(–13.38)	(–13.87)	(–0.11)	(0.25)	(2.81)	(1.75)	(5.49)	(4.35)
OptQSpr/ FutQSpr	–0.003***	–0.034***	0.503**	4.488**	–0.026***	–0.248***	–0.045	–47.240	0.108***	101.680***	0.001	–0.582
	(–7.18)	(–8.42)	(2.44)	(2.19)	(–6.65)	(–6.31)	(–1.48)	(–1.58)	(5.21)	(4.88)	(0.33)	(–0.10)
Skew	–0.488***	–0.432***	0.219	0.390*	–0.230***	–0.159***	0.389***	0.389***	–0.075***	–0.060***	–0.181***	–0.271***
	(–55.02)	(–49.35)	(0.99)	(1.77)	(–9.15)	(–7.30)	(4.85)	(5.04)	(–4.63)	(–3.69)	(–9.39)	(–14.50)
Vola	–0.768***	–0.935***	–1.014	–1.065*	–0.094	–0.594***	0.937*	0.759	–0.659***	–0.767***	–0.363***	–0.825***
	(–39.53)	(–59.18)	(–1.55)	(–1.64)	(–0.42)	(–2.80)	(1.76)	(1.43)	(–8.07)	(–10.81)	(–4.38)	(–12.88)
N	97	439	97	439	97	467	97	467	97	447	97	447
Adj. R^2	0.33	0.34	0.06	0.07	0.14	0.13	0.05	0.05	0.06	0.06	0.07	0.06

Table 4.7: Speculation in commodity derivative markets (continued)

Panel B	Crude oil		Natural gas		Gold		Silver		Corn		Soybeans	
	Model 1: *WorkingT*	Model 2: *R*	Model 1: *WorkingT*	Model 2: *R*	Model 1: *WorkingT*	Model 2: *R*	Model 1: *WorkingT*	Model 2: *R*	Model 1: *WorkingT*	Model 2: *R*	Model 1: *WorkingT*	Model 2: *R*
Intercept	0.288***	2.595***	0.130	0.008	0.497***	2.214***	0.229**	–0.526	–0.023	1.339***	–0.166***	0.496***
	(25.71)	(53.51)	(1.27)	(0.01)	(10.04)	(27.59)	(2.38)	(–1.37)	(–0.82)	(12.72)	(–4.35)	(8.38)
Metric	0.102***	–10.424***	0.024	5.746***	0.295***	–0.777***	0.161**	–0.354**	0.081***	–0.683***	0.168***	–0.513***
	(9.31)	(–26.77)	(0.40)	(3.79)	(6.24)	(–7.40)	(2.58)	(–2.36)	(4.30)	(–5.70)	(5.01)	(–6.23)
OptVol/FutVol	–2.132***	–2.544***	–3.331	2.429	–0.673*	–0.975***	4.675	4.946	–1.140***	–1.300***	0.280***	0.337***
	(–11.89)	(–14.30)	(–0.36)	(0.27)	(–1.93)	(–2.80)	(1.37)	(1.45)	(–6.22)	(–7.17)	(2.66)	(3.25)
OptOI/FutOI	–0.036***	–0.668***	–0.59**	–5.684**	–0.874***	–0.875***	–2.261***	–1.042*	0.596***	0.309***	0.243***	0.185***
	(–6.72)	(–12.80)	(–2.58)	(–2.54)	(–17.73)	(–17.87)	(–3.25)	(–1.80)	(5.62)	(2.70)	(4.73)	(3.54)
OptQSpr/FutQSpr	–0.002***	–0.002***	0.472***	0.425**	–0.014***	–12.890***	0.006	–0.043	0.130***	112.68***	–0.002	–3.220
	(–4.95)	(–6.01)	(2.67)	(2.44)	(–4.09)	(–3.75)	(0.24)	(0.00)	(5.92)	(5.15)	(–0.32)	(–0.57)
Skew	–0.331***	–0.281***	–0.008	0.076	–0.272***	–0.189***	0.225***	0.270***	–0.089***	–0.051***	–0.147***	–0.208***
	(–41.65)	(–35.96)	(–0.04)	(0.42)	(–12.33)	(–9.97)	(3.09)	(3.83)	(–5.18)	(–3.00)	(–8.03)	(–11.69)
Vola	–0.824***	–0.952***	–0.707	–0.810	0.685***	0.095	1.972***	1.617***	–1.120***	–1.303***	–0.480***	–0.723***
	(–47.26)	(–67.42)	(–1.28)	(–1.49)	(3.52)	(0.51)	(4.07)	(3.34)	(–13.03)	(–17.49)	(–6.08)	(–11.90)
N	97	439	97	439	97	467	97	467	97	447	97	447
Adj. R^2	0.31	0.33	0.04	0.07	0.18	0.18	0.05	0.05	0.10	0.10	0.04	0.04

Chapter 5: Liquidity and Earnings in Event Studies: Does Data Granularity Matter?

5.1 Overview

In the last century, data availability has improved significantly. The granularity of microstructure liquidity data have moved through various stages, from yearly, to quarterly, to monthly, to daily prices and trade volumes and following on to detailed transaction and quote data at the nanosecond level. However, it is unclear whether the increase in data granularity is necessary to increase understanding of financial markets. Recently, a number of concerns have emerged, including the growth in size of microstructure data and computational burdens, the expense of intraday data and the use of LF liquidity proxies in microstructure, asset pricing and corporate finance research. Specific concerns relate to whether such proxies capture transaction costs or marginal costs per dollar volume.[25] Consequently, Goyenko, Holden and Trzcinka (2009) and Fong, Holden and Trzcinka (2017) concluded that a subset of LF liquidity measures estimated using annual, monthly and daily equity data can proxy for HF benchmarks calculated using intraday data.[26] These findings show that LF data have some advantages compared to those that are HF. For example, reduced processing times and cost savings for researchers (particularly in countries where data are unavailable or expensive) and the ability for liquidity research to be conducted in a wider range of finance and accounting topics, using larger samples and across longer time periods.

Although the results of these prior studies are largely robust across countries and time, this study directly tested the findings of Goyenko et al. (2009) and Fong et al. (2017) and examined whether LF liquidity proxies (if any) can proxy for HF benchmarks when there is information asymmetry in the market, specifically around Australian earnings announcements. The focus is on earnings event studies for two reasons. First, the event study methodology is a widely used framework. For example, there are more than 1700

[25] To illustrate the magnitude of the computational savings achieved by using LF liquidity measures, Fong et al. (2017) reported that the ratio of HF to LF data points was 42 times in 1996 for their multi-country sample, which grew exponentially to 12,616 in 2014.

[26] The ability of LF liquidity measures to capture HF benchmarks has also been examined in commodity futures (e.g., Marshall et al., 2012) and corporate bonds (e.g., Schestag et al., 2016).

papers today in SSRN with the phrase 'event study' in the title, abstract or keywords. Second, consistent with the findings of Collin-Dufresne and Fos (2015), it is anticipated that the performance of LF measures as HF proxies may not hold during periods of stress (e.g., periods of information asymmetry). Therefore, this analysis will indicate whether the granularity of data matter and if computational and cost savings are available in event studies research. These findings will also be interesting to market makers, investors, academics and regulatory bodies when conducting their financial activities.

Liquidity measures can be broadly split into two groups: 1) percent–cost measures that capture transaction costs and 2) cost-per-dollar measures that capture the marginal transaction costs per unit of dollar volume. The correlation between low and HF liquidity measures is compared in an event study framework around Australian earnings announcements. At the daily level for the percent–cost measures, the performance of two LF measures are compared—high–low (HL) (Corwin & Schultz, 2012) and closing-percent-quoted-spread (CPQS) (Chung & Zhang, 2014)—to four HF benchmarks: quoted spreads, effective spreads, realised spreads and price impact. In addition, for the cost-per-dollar measures, four LF measures were compared—Amihud (2002), Amivest, HLI and closing-percent-quoted-spread-impact (CPQSI) (Goyenko et al., 2009)—to one HF benchmark: lambda (Hasbrouck, 2009).

For the percent–cost measures, it was found that the CPQS is the best proxy for quoted spreads, effective spreads, price impact and realised spreads. This result is largely consistent using cross-sectional, portfolio and individual time-series correlations. Root-Mean-Squared-Error (RMSE) analysis was used to capture the level of the HF benchmarks. This was conducted during the full sample, event, benchmark, pre-announcement and post-announcement periods around the earnings announcement and sub-sampling on earnings surprises, stock size and macro-economic and institutional changes in the Australian market.

In contrast, it was found that HLI and CPQSI tend to be the most consistent cost-per-dollar proxies of lambda using cross-sectional, portfolio and individual time-series correlations. However, the performance of these proxies weakens in the pre and post-announcement periods around the earnings announcement. Further, the Amihud measure is the best proxy for capturing the level of lambda. The results of this study indicate that caution should be taken when using the LF cost-per-dollar measures to proxy for HF lambda during periods of increased information asymmetry, especially if these measures are used in time series analysis.

In particular, evidence was found that indicated that LF measures can proxy for HF benchmarks during periods of information asymmetry for the percent–cost measures. This conclusion is at odds with the findings by Collin-Dufresne and Fos (2015). A plausible explanation for why LF measures can proxy for HF measures around earnings announcements is that, as the lead time before an informed trader's private information becomes public decreases, insiders have less ability to be strategic. If insiders cannot strategically time their trades (e.g., during times of high liquidity), then liquidity measures (either LF or HF) can capture informed trading or behave consistently with microstructure theory (Ahern, 2017).

These findings indicate that various parties can use the CPQS to examine percent–cost liquidity in event studies analysis (e.g., liquidity providers in setting spreads or regulators examining informed trading). These findings are robust during times when private information trading is of greater interest; that is, times of increased information asymmetry, during positive and negative earnings, for large stocks and more recent time periods. In addition, these results provide initial evidence that LF measures can be used in event studies analysis in markets and during periods when HF data are unavailable, while simultaneously providing cost and computational savings.

This chapter proceeds as follows. Section 5.2 describes the data. Section 5.3 details the methodology used to calculate low and HF liquidity measures, event study framework and correlation and RMSE analysis. Section 5.4 reports the findings and Section 5.5 concludes.

5.2 Data

Trade and quote data and end-of-day data were used to compute the HF and LF measures from the *TRTH* database. Data were obtained from the market quality dashboard that is developed and managed by the CMCRC. This sample consists of 222 firms listed on the

Australian Securities Exchange (ASX) 200 index between 1 January 2006 and 31 December 2016. Due to the inaccuracy of dates on the I/B/E/S database, half-yearly earnings announcements and individual analyst forecasts were obtained from *Bloomberg* (Barber et al., 2013; Creighton et al., 2007; Griffin et al., 2011; Pevzner et al., 2015). Firms were kept in this analysis that were removed from the ASX 200 index and the *TRTH* reference change list was used to account for firm symbol changes during the 10-year sample period.

5.3 Methodology

There are two classes of liquidity measures in the literature: 1) percent–cost measures that capture the costs to complete small transactions and 2) cost-per-dollar measures that capture the additional costs for trading each dollar of volume. Following the methodology by Fong et al. (2017), four HF and two LF percent–cost measures were calculated on a daily basis, in addition to one HF and four LF cost-per-dollar measures.

5.3.1 Percent–cost Measures

Four daily HF percent–cost measures were calculated using trade and quote data. The first HF measure is quoted spreads, which capture the round trip transaction costs of small transactions:

$$Quoted\ Spread_s = \frac{Ask_s - Bid_s}{(Bid_s + Ask_s)/2}, \tag{5.1}$$

where Bid_s and Ask_s are the best quotes at time s. A daily quoted spread measure was obtained for stock i on day t by taking the time-weighted average of $Quoted\ Spread_s$ during day t.

The second HF measure is effective spreads, which captures the round trip costs of a liquidity demanding trade:

$$Effective\ Spread_k = 2D_k \frac{P_k - ((Bid_k + Ask_k)/2)}{(Bid_k + Ask_k)/2}, \tag{5.2}$$

where P_k is the price of trade k, Bid_k (Ask_k) is the bid (ask) price at the time of trade k and D_k is a buy or sell indicator of trade k equal to +1 for buy volume and equal to –1 for sell volume.[27] A daily effective spread measure was obtained for stock i on day t by taking the volume-weighted average of $Effective\,Spread_k$ during day t. A similar process is applied to estimate a daily price impact and daily relative spread (formulae detailed below).

The third HF measure is price impact, which captures the permanent price change following a transaction:

$$Price\ Impact_k = 2D_k \frac{((Bid_{k+5}+Ask\text{–}_{k+5})/2) - ((Bid_k+Ask_k)/2)}{(Bid_k+Ask_k)/2}, \tag{5.3}$$

where $k+5$ indicates the mid-quote five minutes after trade k.

The fourth HF measure is realised spreads, which captures the costs of a liquidity demanding trade net of the price impact of the trade (i.e., effective spreads = price impact + realised spread):

$$Realised\ Spread_k = 2D_k \frac{P_k-(Bid_{k+5}+Ask_{k+5})/2}{(Bid_k+Ask_k)/2}. \tag{5.4}$$

These four HF percent–cost measures are compared with two LF liquidity proxies that are calculated using daily data. The first LF proxy is the HL measure developed by Corwin and Schultz (2012), which estimates a daily bid-ask spread using daily high and low prices:

$$HL_t = \frac{2(e^{\alpha_t}-1)}{1+e^{\alpha_t}}, \tag{5.5}$$

where $\alpha_t = \frac{\sqrt{2\beta_t}-\sqrt{\beta_t}}{3-2\sqrt{2}} - \sqrt{\frac{\gamma_t}{3-2\sqrt{2}}}$, β_t is the sum of the squared ratio of daily high to daily low prices expressed in log form across day $t-1$ and day t and γ_t is the squared ratio of the two-day high price to the two-day low price expressed in log form. Similar to Corwin and Schultz (2012), negative spread estimates are set to zero.

[27] The methods by Lee and Ready (1991) are followed to determine the buy or sell indicator.

The second LF proxy is the CPQS developed by Chung and Zhang (2014), which also estimates a daily bid-ask spread. However this uses closing ask and closing bid prices:

$$CPQS_t = \frac{Closing\ Ask_t - Closing\ Bid_t}{(Closing\ Ask_t + Closing\ Bid_t)/2}, \tag{5.6}$$

where $Closing\ Ask_t$ ($Closing\ Bid_t$) is the closing ask (bid) price on day t. Similar to Fong et al. (2017), if intraday quote data are missing, the CPQS is set to missing. In addition, valid values are required for both HL and CPQS, otherwise both measures are set to missing.

5.3.2 Cost-per-dollar Measures

One daily HF cost-per-dollar measure is calculated using trade and quote data. In a similar vein to Hasbrouck (2009), lambda is calculated, which captures the cost for demanding liquidity during a five-minute interval (n):

$$r_n = \lambda S_n + \varepsilon_n, \tag{5.7}$$

where r_n is the return for period n, $S_n = \sum D_k \sqrt{|v_{kn}|}$, where v_{kn} is the signed dollar volume of trade k in period, n, ε_n is the error term and λ is lambda calculated for stock i on day t.

Lambda is compared with four LF proxies. The first LF proxy is the Amihud (2002) illiquidity measure, which captures the daily price change resulting from each dollar of volume traded:

$$Amihud_t = \frac{|r_t|}{\$Vol_t}, \tag{5.8}$$

where r_t is the stock return on day t and $\$Vol_t$ is the dollar volume on day t. Similar to prior studies, the Amihud (2002) measure is only calculated on positive volume days.

The second LF proxy is the Amivest measure, which captures the price impact of dollar volume:

$$Amivest_t = \frac{\$Vol_t}{|r_t|}, \tag{5.9}$$

The Amivest measure is calculated on all days in which the daily return is not equal to zero.

The method by Fong et al. (2017) is used to estimate the third and fourth LF proxies, which are the HLI and CPQSI. Both of these proxies are the HL and CPQS measures divided by $\$Vol_t$. These proxies capture the response of daily bid-ask spreads to the dollar volume traded:

$$HLI_t = \frac{(HL_t)}{\$Vol_t}, \tag{5.10}$$

$$CPQSI_t = \frac{CPQS_t}{\$Vol_t}. \tag{5.11}$$

All cost-per-dollar LF proxies are multiplied by 1,000, except Amivest, which is divided by 1,000,000, and HLI and CPQSI, which is multiplied by 1,000,000. In addition, the following filters are applied for both the percent–cost and cost-per-dollar measures:

1) stock months with less than 5 positive volume days and more than 11 zero-return days are excluded
2) LF proxies are winsorized at the 1 percent and 99 percent levels
3) LF proxy is set to missing if it is not possible to compute the HF measure.

All percent–cost HF and LF measures are proxies of spreads. The spread has several components, including an adverse selection component, which captures information asymmetry between liquidity providers and traders. The adverse selection component increases in the presence of informed traders (Glosten & Milgrom, 1985). Empirically, Krinsky and Lee (1996) found that adverse selection costs (or information asymmetry) are heightened around earnings announcements. As a result, positive increases are expected in both HF and LF measures prior to the earnings announcement (due to adverse selection costs). Therefore, a positive correlation between both HF and LF measures is expected.

Given that earnings news increases information asymmetry between investors, this means a widening of the spread (via an increase in adverse selection costs) and an increased spread results in an increase in illiquidity. The cost-per-dollar measures can be considered measures of illiquidity. Therefore, positive increases in lambda and positive increases in

Amihud, HLI and CPQSI are expected prior to the earnings announcement. This is due to increased levels of illiquidity. As such, a positive correlation is anticipated between these LF measures and lambda. However, positive changes in lambda and negative changes in Amivest are expected prior to the earnings announcement. This is because Amivest is the reverse of Amihud (e.g., in Table 5.6, Fong et al. (2017) find a negative relation between lambda and Amivest). As a result, a negative correlation between Amivest and lambda is expected.

5.3.3 Earnings Event Study Framework

The HF and LF measures are calculated for each stock day during the event window around the earnings announcement. Similar to prior studies (e.g., Barraclough et al., 2013; Khan & Lu, 2013; Hao, 2016; Patel & Michayluk, 2016b) the following event study windows were created around the earnings announcement date [0]: 1) event period [–60,+5], 2) benchmark period [–60,–11], 3) pre-announcement period [–5,0] and 4) post-announcement period [0,+5].

In addition to the event period, this research also separately examines the correlation between LF and HF measures during the pre and post-announcement periods. This is because information asymmetry has been found to be heightened during such periods around earnings announcements (Krinsky & Lee, 1996) compared to normal trading periods (e.g., the benchmark period).

5.3.4 Correlation and Root-Mean-Squared-Error Analysis

This research examines whether LF proxies can proxy for HF measures by calculating the correlation between the LF and HF measures during the event, benchmark, pre-announcement and post-announcement windows. Following Fong et al. (2017), four different tests were conducted to examine the performance of LF to HF measures, including computing. These were:

1) average cross-sectional correlations
2) portfolio time-series correlations
3) individual stock time-series correlations
4) RMSE analysis.

Applying a similar two-step approach to Fama and MacBeth (1973), the average cross-sectional correlations of each LF proxy were estimated with each HF measure. First, for each combination of the LF and HF measures, the cross-sectional correlation was estimated across all firms for each day t. Second, the average correlation was calculated during each event study window.

For each LF–HF combination, the portfolio time-series correlations were calculated using the average value of the LF and HF measures on day t across all stocks and then the time series correlation of the LF–HF combination during each event study window. In addition, for each stock i, the individual stock time-series correlations were calculated for each LF–HF combination during the event study windows. The average was then reported.

When reporting the average cross-sectional portfolio time series and individual stock time-series correlations for each HF measure a solid box is placed around the highest LF–HF correlation, a dashed box is placed around each LF–HF correlation that is not statistically different from the highest LF–HF correlation and an LF–HF correlation is highlighted in bold text if the correlation is statistically different from zero at a five percent level of significance.

To determine whether a dashed box is placed around a cross-sectional correlation, the correlation of one LF–HF combination is regressed on the correlation of another LF–HF combination and the Newey-West corrected standard errors are used with four lags. Using a Fisher Z-test, 95 percent confidence intervals are generated to determine whether a dashed box is placed around the time series correlations.

In addition, the RMSE for each LF–HF combination for each day t and the average RMSE is reported across each event study window. The average RMSE indicates the similarity of the levels of the LF proxy and HF measure. A lower average RMSE indicates that a particular LF proxy is better at capturing the level of the HF measure. For each HF measure, a solid box is placed around the smallest average RMSE, a dashed box is placed around average RMSEs that are not statistically different from the lowest average RMSE and an average RMSE is highlighted in bold text if the RMSE can predict the level of the HF measure at a five percent level of significance. A paired *t*-test is used to determine whether a dashed box is placed around an average RMSE value. A *U*-statistic is used to determine whether an average RMSE is reported in bold text (Theil, 1966).

5.4 Results

This section examines whether LF liquidity measures can proxy for HF benchmarks during the full sample and periods of varying information asymmetry around earnings announcements.

5.4.1 Descriptive Statistics

Table 5.1 reports descriptive statistics for each of the HF and LF measures during the full sample (Panel A), event period (Panel B), benchmark period (Panel C), pre-announcement period (Panel D) and post-announcement period (Panel E). The full sample captures over 390,000 observations of daily HF and LF liquidity measures, of which 21,749 are estimated during the earnings announcement event period. For the full sample, the percent–cost HF benchmarks have a mean range of 0.012 (price impact %) to 0.032 (quoted spread %) and the LF counterparts range from 0.007 (HL) to 0.011 (CPQS). The cost-per-dollar LF proxies have mean values that range between 0.012 (CPQSI) and 7.629 (Amihud) when compared to the average value of lambda (–0.004). Comprehensively, the mean values of each LF and HF variable reported in Table 5.1 are similar to those reported in Fong et al. (2017). Generally, the mean and median values for all key variables are similar. The exceptions are Amihud, Amivest and HLI, which are left skewed.

< Table 5.1 here >

Figure 5.1 plots the behaviour of the LF/HF measures during the period [–11,+5] around the earnings announcement. Panel A captures the percent–cost measures and Panel B

captures the cost-per-dollar measures. This period [–11,+5] corresponds with a period of heightened information asymmetry around the earnings announcement. Panels A and B show that most liquidity measures remain relatively constant during the event period, with a slight increase around the announcement date (day 0). In contrast, the lambda variable is particularly volatile over this period, consistent with the findings reported in Fong et al. (2017).

< Figure 5.1 here >

5.4.2 Correlations and Root-Mean-Squared-Error Analysis

This research separately examines whether percent–cost and cost-per-dollar LF measures can proxy for HF measures using both correlation and RMSE analysis during the full sample period (see Table 5.2), event period [–60,+5] (see Table 5.3), benchmark period [–60,–11] (see Table 5.4), pre-announcement period [–5,0] (see Table 5.5) and post-announcement period [0,+5] (see Table 5.6). The focus is to examine whether the general findings in Goyenko et al. (2009) and Fong et al. (2017) are applicable during periods of information asymmetry between investors. Three types of correlations are estimated (average cross-sectional, portfolio time series, individual time series) to assess the best LF proxy for each HF measure. The average RMSE is reported, which indicates whether a particular proxy captures the level of a benchmark. A solid box indicates the dominant measure (in terms of correlation or RMSE) for that group and a dashed box indicates if the correlation or RMSE is statistically similar to the dominant correlation or RMSE. Bold text indicates if the correlation is statistically different from zero or if the RMSE can statistically predict the level of the HF measure.

5.4.2.1 Percent–cost Measures

Table 5.2 Panel A shows that, although HL is the dominant proxy for all four HF measures during the full sample period, CPQS is statistically similar. Further, the reported correlations are similar in level to Fong et al. (2017). For example, Fong et al. (2017)

reported cross-sectional correlations of 0.343 and 0.728 between the CPQS and HF measures. This research respectively reports correlations that range from 0.348 to 0.552.

< Table 5.2 here >

The results are less consistent for the portfolio time-series LF–HF correlations reported in Panel B, in which the LF measures are equally dominant for quoted and effective spreads. However, the CPQS is the dominant proxy for price impact and HL is the dominant proxy for realised spreads. Panel C reports the individual time-series correlations and shows that CPQS is the dominant measure.

The average RMSE for each LF–HF combination is reported in Panel D. A lower RMSE indicates that an LF measure is better at capturing the level of the HF measure. Both the CPQS and HL do a statistically similar job in capturing the level of the HF measures.

Both the HL and CPQS are adequate proxies for their HF counterparts. The CPQS is the best, or joint best, LF proxy for the HF measures in 15 out of the 16 correlation and RMSE tests across Panels A to D. In contrast, the HL measure is the dominant, or equivalent best, proxy in 11 of the 16 tests. These findings are similar to Fong et al. (2017), who also reported that the CPQS is the dominant or equivalent in 14 out of 16 tests. However, their results for the HL proxy are weaker compared to the findings of this study.

Table 5.3 reported correlations and RMSEs for the LF and HF measures during the [–60,+5] period around the earnings announcement. Similar to the full sample findings in Table 5.4, it was found that the CPQS is the dominant or equivalent best proxy for the HF benchmarks in 16 out of 16 tests across Panels A to D and the magnitudes of the correlations and RMSEs are similar.

< Table 5.3 here >

< Table 5.4 here >

A similar trend is observed in Table 5.5 (pre-announcement period) and Table 5.6 (post-announcement period), in which CPQS is the best or equivalent best proxy for the quoted spread, effective spread, price impact and realised spread measures in terms of correlations and RMSE. Similar to a Ball and Brown (1968) event study framework, asterisks are reported if the correlations and RMSEs in Table 5.5 and Table 5.6 are significantly different at the five percent level from the respective correlations and RMSEs in Table 5.4 (i.e., the benchmark period). In both Tables 5.5 and 5.6, almost all the CPQS portfolio time-series or individual time-series correlations are statistically different to those reported in the benchmark period. This finding suggests that heightened levels of information asymmetry around the earnings announcement affect the ability of the CPQS to proxy for the HF counterparts in the time series.[28] The cross-sectional correlations and RMSEs are statistically similar across Tables 5.4 to 5.6.

< Table 5.5 here >

< Table 5.6 here >

Overall, it was found that the CPQS is the dominant percent–cost proxy of the HF benchmarks using different methods for estimating correlations, across periods of differing levels of information asymmetry and for different HF measures.

5.4.2.2 Cost-per-dollar Measures

The full sample results reported in Table 5.2 indicate that the HLI and CPQSI are the best proxies for lambda in terms of cross-sectional, portfolio time-series and individual time-series correlations. Universally based on correlations, Fong et al. (2017) also found that CPQSI is the dominant LF proxy. In contrast to the percent–cost correlations, the cost-per-dollar correlations are smaller in magnitude. For example, the cross-sectional correlations (Panel A) are at approximately the 30 percent level for the cost-per-dollar measures (except for Amivest) compared to approximately 60 percent for the percent–cost measures.

In contrast to Panels A to C, which indicate that HLI and CPSQI are the dominant proxies for lambda, Panel D indicates that Amihud is the dominant proxy for capturing the level of lambda (RMSE of close to zero). Holistically, throughout the analysis, Amivest is the weakest proxy for lambda in terms of both correlations and RMSEs.

[28] Most of the HL correlations or RMSEs estimated during the pre (12/16) and post-announcement (9/16) periods significantly differ from those reported during the benchmark period.

Overall, the full sample cost-per-dollar LF proxies produce mixed conclusions. If interested in using LF proxies to capture correlations, then the most consistent proxies of the four that are considered are the HLI and CPQSI. However, if interested in capturing the level of lambda, then the Amihud measure should be considered. The inconclusive results are consistent with the noise embedded in the lambda measure (Fong et al., 2017).

Other than the magnitude of the cost-per-dollar portfolio time-series correlations, the main conclusions in Table 5.2 also hold in Table 5.3 and Table 5.4; that is, during the event period [–60,+5] and benchmark period [–60,–11] relative to the earnings announcement.

Again, Table 5.5 and Table 5.6 compare the correlations and RMSEs during the pre and post-announcement periods to their equivalent values during the benchmark period (see Table 5.4). Using cross-sectional correlations (RMSEs), it was found that, similar to Table 5.4, the HLI and CPQSI (Amihud) measures are the dominant proxies for lambda in Tables 5.5 and 5.6. However, no statistical difference was reported between these values. The increased levels of information asymmetry in the 10 days surrounding the earnings announcement show that, using portfolio time-series (individual time series) correlations, Amivest (HLI) is the dominant (dominant or equivalent) proxy for lambda. In addition, it was observed that the portfolio and individual time-series correlations for HLI and CPQSI are significantly weaker from the benchmark to the pre and post-announcement periods, as indicated by asterisks in Table 5.5 Table 5.6.

During periods of information asymmetry around earnings announcements, mixed findings were observed as to the dominant proxy for lambda, in particular in the time series in which the magnitude of the correlations is considerably closer to zero. Based on these findings, researchers should exercise caution when using the cost-per-dollar LF proxies used in this study to capture lambda during periods of information asymmetry, especially if these measures are used in a time series manner.

Due to the strategic trading of informed traders, Collin-Dufresne and Fos (2015) found that liquidity measures do not always behave consistently with microstructure theory. In contrast to the predictions in this study—in particular for percent–cost measures—it was reported that LF measures can consistently proxy for HF benchmarks, even during periods of information asymmetry. Such a result is possible because, in the period before earnings announcements, the ability for strategic informed trading decreases. As a result, informed traders cannot time their trades during periods of lower transaction costs. Therefore, liquidity measures (LF or HF) can capture the presence of informed trading and behave consistently with microstructure theory (Ahern, 2017).

5.4.2.3 Robustness

In this section, the robustness of the main findings are examined. Appendix A reports the performance of LF versus HF measures during the event period [–60,+5] and sub-sampling on positive, negative and no-earnings surprises, large and small stocks, pre and post-global financial crisis (GFC) and before and after the introduction of the Chi-X stock exchange into the Australian market. Appendix B is similar to Appendix A; however, it reports the performance of LF and HF measures during the pre-announcement period [–5,0].

Similar to Bernile et al. (2016), earnings surprises are determined using the difference between the announcement value and the survey median earnings analyst value, normalised by the survey standard deviation. Large (small) stocks are defined using the top (bottom) tercile of stocks based on market capitalisation. The method devised by Beber and Pagano (2013) was followed and the pre-GFC period was defined as prior to January 2008. The post-GFC period was defined as following June 2009. The pre-Chi-X and post-Chi-X periods were defined as either side of 30 October 2011.

5.4.2.4 Percent–cost Measures Robustness

Positive, negative and no-earnings surprises (see Tables 5.7 to 5.9 and Tables 5.16 to 5.18: In general, for positive and negative earnings surprises, event and pre-announcement periods and using both correlations and RMSEs, CPQS is the dominant proxy for the HF measures. This is consistent with the main analysis. In contrast, the relative dominance of the HL measure increases for earnings news, which is not a surprise to the market.

Large and small stocks (see Tables 5.10 to 5.11 and Tables 5.19 to 5.20): Sub-sampling by size yields different outcomes for large and small stocks. For large stocks, similar findings are reported to the main findings of this study. That is, CPQS is the better measure in terms of correlations and level when proxying for quoted spreads, effective spreads, price impact and realised spreads. The relative performance of the HL proxy is improved for smaller stocks. By definition, the HL measure is subject to outliers or smaller firms, which are subject to more volatile prices. These size sub-sample results are consistent across the event and pre-announcement periods.

Pre and post-GFC (see Tables 5.12 to 5.13 and Tables 5.21 to 5.22): Generally, similar conclusions are drawn between the pre-GFC (post-GFC).

Before and after the introduction of the Chi-X stock exchange (see Tables 5.14 to 5.15 and Tables 5.23 to 5.24): Overall, the CPQS is a better proxy for HF measures before and after the introduction of the Chi-X stock exchange. They are relatively robust between different periods of information asymmetry (i.e., event versus pre-announcement periods).

Cost-per-dollar measures: Relative to the percent–cost measures, the performance of the cost-per-dollar measures is weaker in terms of the magnitude of LF and HF correlations and as to the dominant LF proxy. For the following sub-samples and for both the event and pre-announcement periods, positive earnings surprises, no-earnings surprises, small stocks, pre-GFC, post-GFC and post-Chi-X, either the HLI or CPQSI has a solid or dashed box, indicating that these two proxies are the more consistent proxy for lambda (i.e., they have the highest correlation with lambda). In all sub-samples, the Amihud measure best captures the level of lambda, which is consistent with the main analysis of this research.

5.5 Summary

Market microstructure data availability has significantly improved and it is now possible to estimate liquidity measures at the nanosecond level. However, this level of data is unavailable in all markets and time periods and there is a significant cost and computational burden of HF data. Goyenko et al. (2009) and Fong et al. (2017) show that various LF liquidity measures can proxy for HF benchmarks and that the results are robust across countries and time. However, liquidity measures do not always behave as expected during periods of information asymmetry (Collin-Dufresne & Fos, 2015). Drawing from Ball and Brown (1968), an event study methodology is used to investigate whether the LF measures of liquidity can proxy for HF measures around earnings announcements (i.e., periods of information asymmetry).

This research found that the CPQS is the best proxy for the percent–cost HF benchmarks. This result is largely robust using different correlation methods, using the RMSE to capture the level of the HF benchmarks, across different periods with differing levels of information asymmetry and using sub-sampling on earnings surprises, stock size and around macro-economic and institutional changes in the Australian market.

Further, the CPQS can be used as a consistent LF percent–cost proxy to measure liquidity during positive and negative earnings surprises, in large stocks and in recent time periods. That is, during times of interest to a number of individuals. For example, traders are more likely to conduct informed trading around earnings surprises than non-surprises, regulators can use earnings surprises as a signal to investigate cases of informed trading and larger stocks are of greater interest to a wider range of market participants.

In contrast, using cross-sectional, portfolio and individual time-series correlations, the most consistent LF proxies are the HLI and CPQSI. However, the performance of these proxies weakens in the pre and post-announcement periods around the earnings announcement. Further, the Amihud measure is the best proxy for capturing the level of lambda. In particular, in time series analysis, it is suggested that caution should be exercised when using the cost-per-dollar measures to proxy for lambda during periods of increased information asymmetry.

In contrast to predictions based on the findings of Collin-Dufresne and Fos (2015), this research found that LF measures (in particular CPQS) can proxy for HF benchmarks during periods of information asymmetry. Ahern (2017) suggested that the ability for informed traders to behave strategically diminishes in the days around the earnings announcement. As a result, liquidity measures may behave consistently with microstructure theory. This means that LF can proxy for HF measures, as they capture information asymmetry prior to the earnings announcement.

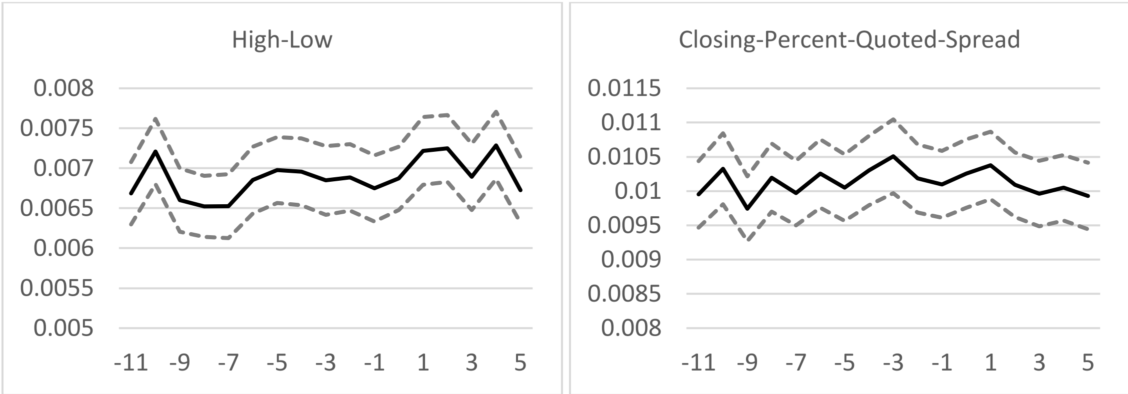
There are several directions suggested for future research. First, this paper is agnostic on the choice of HF and LF proxies and uses the same set of variables as Fong et al. (2017). There is an opportunity to widen the scope of the liquidity measures; in particular, additional HF cost-per-dollar benchmarks for robustness. Second, this research focuses on scheduled earnings announcements as a proxy for periods of information asymmetry. Future research may consider examining unscheduled news announcements, which allow for a cleaner measurement of information asymmetry.

Figure 5.1: Daily plots of LF and HF liquidity measures [–11, +5]

This figure depicts the daily correlations between LF and HF measures during the period [–11,+5] around the earnings announcement. Panel A illustrates the percent–cost measures and Panel B illustrates the cost-per-dollar measures. The LF percent–cost measures include HL spread (Corwin & Schultz, 2012) and CPQS (Chung & Zhang, 2014). The HF percent–cost measures include quoted spreads (%), effective spreads (%), price impact (%) and realised spreads (%). The LF cost-per-dollar measures include illiquidity (Amihud, 2002), Amivest, high-low-spread-impact and CPQSI. The cost-per-dollar measures include lambda (Hasbrouck, 2009). The sample comprises 222 firms during the period from 1 January 2006 to 31 December 2016.

Panel A: Percent-cost measures

LF measures



HF measures

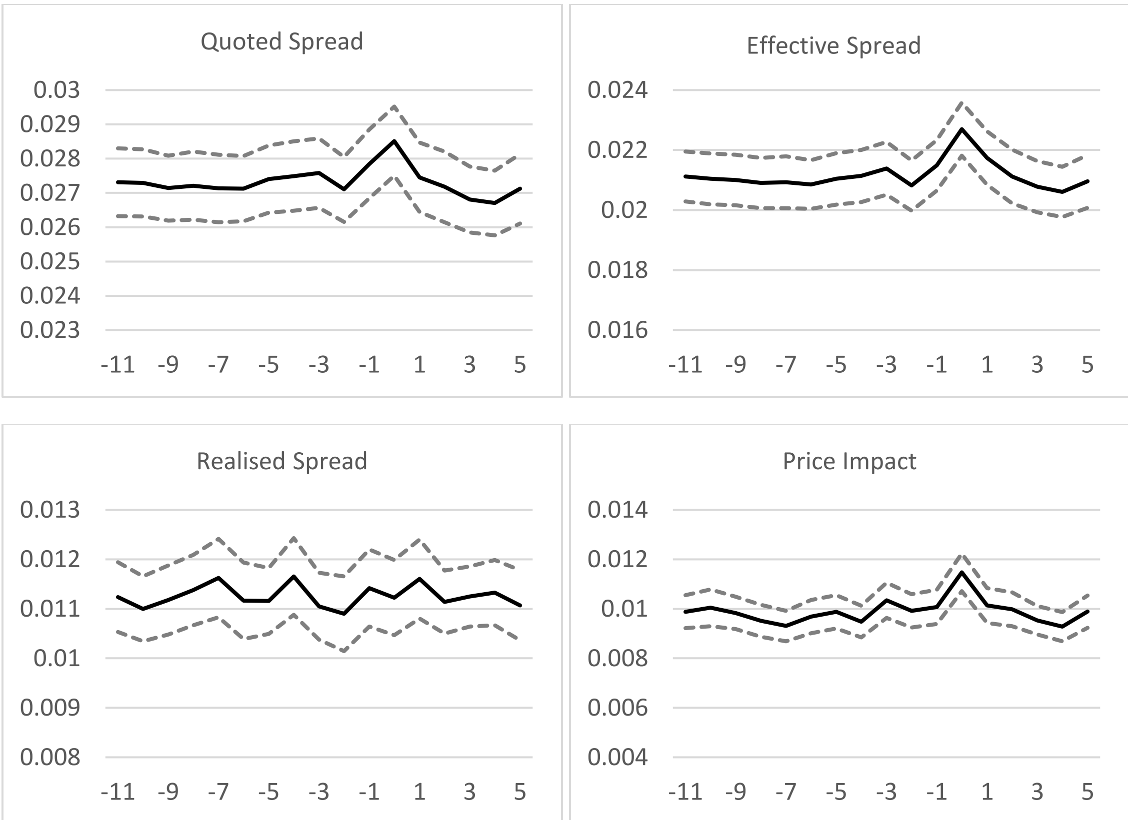
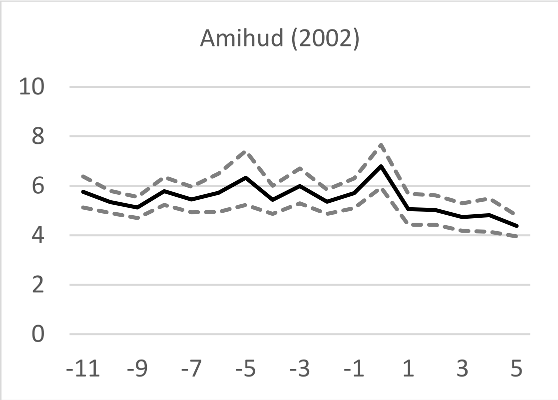
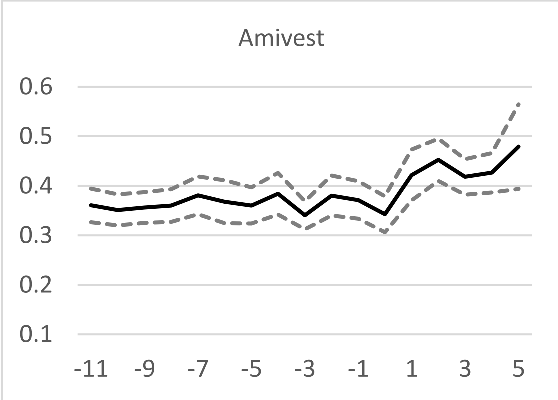


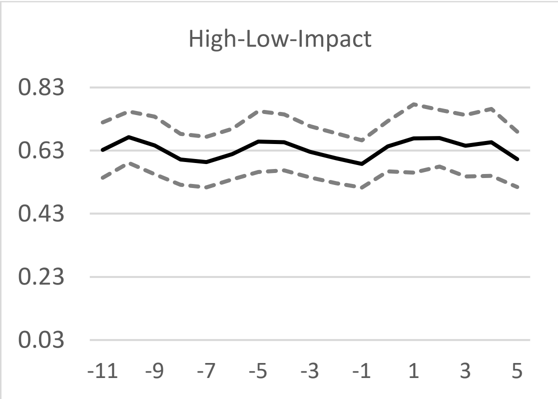
Figure 1: Daily plots of LF and HF liquidity measures [-11,+5] (continued)

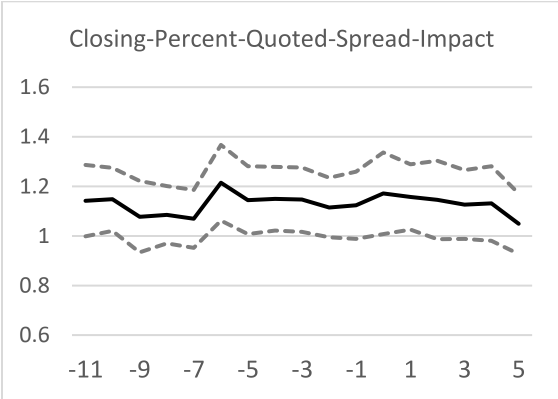
Panel B: Cost-per-dollar measures

LF measures









HF measures

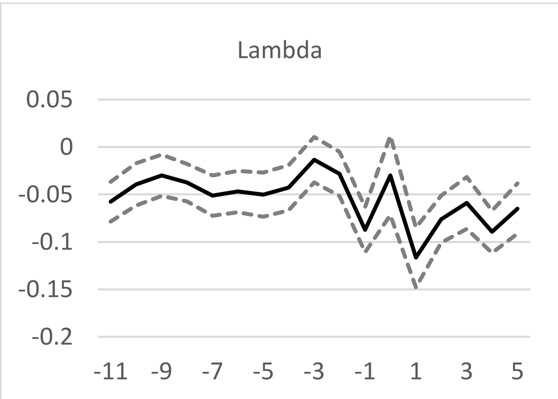


Table 5.1: Descriptive statistics of low- and high-frequency measures

This table reports descriptive statistics for LF and HF measures. The LF percent-cost measures include: high-low (HI) spread (Corwin and Schultz, 2012) and closing-percent-quoted-spread (CPQS) (Chung and Zhang, 2014). The HF percent-cost measures include: quoted spreads (%), effective spreads (%), price impact (%) and realised spreads (%). The LF cost-per-dollar measures include: illiquidity (Amihud, 2002), Amivest, high-low-spread-impact (HLI) and closing-percent-quoted-spread-impact (CPQSI). The cost-per-dollar measures include: lambda (Hasbrouck, 2009). Panel A covers the full sample period between January 1, 2006 and December 31, 2016. Panel B covers the event period [-60, +5] around the earnings announcement. Panel C covers the benchmark period (-60 to -11), Panel D covers the pre-announcement period [-5, -0] prior to/during the earnings announcement and Panel E covers the post-announcement period [0, +5] during/following the earnings announcement. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016.

	High-Frequency Benchmarks					**Low-Frequency Benchmarks**					
	Percent-cost				**Cost-per-dollar**	**Percent-cost**		**Cost-per-dollar**			
	Quoted spread	Effective spread	Price impact	Realised spread	Lambda	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Full Sample											
N	393,498	394,589	342,887	394,589	394,589	394,589	394,589	393,498	394,589	342,887	342,887
Minimum	0.001	0.000	0.001	0.001	-0.531	0.000	0.000	0.000	0.000	0.000	0.000
Maximum	0.155	0.164	0.120	0.188	0.686	0.126	0.224	337.0	9.085	8.252	1.043
Mean	0.032	0.025	0.012	0.014	-0.004	0.007	0.011	7.629	0.034	0.687	0.012
Median	0.026	0.019	0.007	0.009	-0.004	0.004	0.008	1.795	0.011	0.109	0.003
Std. Dev.	0.033	0.029	0.023	0.024	0.066	0.009	0.011	2.847	0.091	0.020	0.031
Panel B: Event period [-60,+5]											
N	21,749	21,749	21,749	21,749	21,749	21,749	21,749	21,749	21,749	21,749	21,749
Minimum	0.003	0.002	0.001	0.001	-0.793	0.003	0.002	0.035	0.001	0.011	0.009
Maximum	0.144	0.108	0.048	0.107	0.725	0.019	0.039	660.5	3.742	5.508	9.733
Mean	0.029	0.022	0.010	0.013	-0.036	0.007	0.011	8.980	0.411	0.689	1.266
Median	0.024	0.018	0.008	0.009	-0.036	0.007	0.010	1.199	0.244	0.341	0.451
Std. Dev.	0.026	0.020	0.008	0.015	0.183	0.002	0.007	62.691	0.493	0.963	1.876

Table 5.1: Descriptive statistics (continued)

	High-Frequency Benchmarks					Low-Frequency Benchmarks					
	Percent-cost				Cost-per-dollar	Percent-cost		Cost-per-dollar			
	Quoted spread	Effective spread	Price impact	Realised spread	Lambda	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel C: Benchmark period [-60,-11]											
N Obs.	23,695	23,695	23,695	23,695	23,695	23,695	23,695	23,695	23,695	23,695	23,695
Minimum	0.001	0.001	0.001	0.001	-0.121	0.000	0.000	0.001	0.000	0.000	0.000
Maximum	0.151	0.138	0.052	0.532	0.261	0.101	0.162	185.5	13.628	7.158	1.043
Mean	0.023	0.018	0.008	0.010	-0.003	0.007	0.010	3.055	0.418	0.672	0.012
Median	0.018	0.013	0.005	0.006	-0.004	0.004	0.007	0.688	0.206	0.999	0.003
Std. Dev.	0.023	0.019	0.016	0.018	0.059	0.009	0.011	1.898	0.701	0.215	0.032
Panel D: Pre-announcement period [-5,-0]											
N Obs.	1,624	1,624	1,624	1,624	1,624	1,624	1,624	1,624	1,624	1,624	1,624
Minimum	0.003	0.002	0.000	0.000	-1.121	0.001	0.002	0.029	0.001	0.009	0.007
Maximum	0.131	0.130	0.066	0.069	1.254	0.017	0.041	668.8	5.698	4.262	8.618
Mean	0.028	0.021	0.010	0.011	-0.053	0.007	0.010	5.806	0.371	0.635	1.144
Median	0.022	0.018	0.008	0.008	-0.079	0.007	0.010	0.755	0.366	0.279	0.364
Std. Dev.	0.024	0.021	0.010	0.012	0.272	0.003	0.007	63.417	0.692	0.894	1.685
Panel E: Post-announcement period [0,+5]											
N Obs.	1,948	1,948	1,948	1,948	1,948	1,948	1,948	1,948	1,948	1,948	1,948
Minimum	0.003	0.002	-0.097	0.001	-0.086	0.003	0.002	0.000	0.000	0.000	0.000
Maximum	0.145	0.131	0.059	0.100	0.037	0.035	0.044	513.1	4.178	5.610	0.083
Mean	0.000	0.000	0.010	0.000	-0.073	0.007	0.010	1.174	0.042	0.648	1.130
Median	0.024	0.018	0.008	0.009	-0.001	0.008	0.010	0.117	0.220	0.300	0.005
Std. Dev.	0.025	0.020	0.010	0.014	0.021	0.004	0.007	1.208	0.523	0.983	0.018

Table 5.2: LF vs HF performance – Full sample

This table reports the correlations/RMSEs between LF and HF measures across our full sample period. The LF percent-cost measures include: high-low (HL) spread (Corwin and Schultz, 2012) and closing-percent-quoted-spread (CPQS) (Chung and Zhang, 2014). The HF percent-cost measures include: quoted spreads, effective spreads, price impact and realised spreads. The LF cost-per-dollar measures include: illiquidity (Amihud, 2002), Amivest, high-low-spread-impact (HLI) and closing-percent-quoted-spread-impact (CPQSI). The HF cost-per-dollar measures include: lambda (Hasbrouck, 2009). Panel A reports the average cross-sectional LF-HF correlations, Panel B reports the portfolio time-series LF-HF correlations, Panel C reports the individual stock time-series LF-HF correlations and Panel D reports the average RMSE for each LF-HF combination. In Panels A, B and C, for each HF measure: i) a solid box is placed around the largest LF-HF correlation in each row, ii) a dashed box is placed around each LF-HF correlation which is not statistically different from the largest LF-HF correlation in each row and iii) a LF-HF correlation is highlighted in bold text if the correlation is statistically different from zero at a 5% level of significance. In Panel D, for each HF measure: i) a solid box is placed around the smallest average RMSE in each row, ii) a dashed box is placed around average RMSEs which are not statistically different from the lowest average RMSE in each row and iii) an average RMSE is highlighted in bold text if the RMSE can predict the level of the HF measure at a 5% level of significance. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.627**	**0.552**				
Effective spread	**0.581**	**0.488**				
Price impact	**0.419**	**0.348**				
Realised spread	**0.582**	**0.492**				
Lambda			**0.315**	-0.124	**0.292**	**0.273**
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.538**	**0.5038**				
Effective spread	**0.546**	**0.4960**				
Price impact	**0.282**	**0.3182**				
Realised spread	**0.483**	**0.3988**				
Lambda			**0.380**	**-0.235**	**0.485**	**0.451**
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.240**	**0.364**				
Effective spread	**0.231**	**0.338**				
Price impact	**0.136**	**0.183**				
Realised spread	**0.129**	**0.205**				
Lambda			**0.009**	**-0.030**	**0.040**	**0.046**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.009**	**0.008**				
Effective spread	**0.009**	**0.008**				
Price impact	**0.009**	**0.008**				
Realised spread	**0.009**	**0.008**				
Lambda			**0.000**	5.416	**0.011**	**0.013**

Table 5.3: LF vs HF performance – Earnings event period

This table reports the correlations/RMSEs between LF and HF measures across the event period [-60,+5] around the earnings announcement. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.806**	**0.419**				
Effective spread	**0.823**	**0.407**				
Price impact	**0.700**	**0.370**				
Realised spread	**0.703**	**0.307**				
Lambda			**0.041**	**-0.068**	**0.286**	**0.248**
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.391**	**0.445**				
Effective spread	**0.411**	**0.436**				
Price impact	**0.152**	**0.212**				
Realised spread	**0.341**	**0.314**				
Lambda			**0.136**	**-0.081**	**0.250**	**0.247**
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.210**	**0.314**				
Effective spread	**0.204**	**0.292**				
Price impact	**0.121**	**0.166**				
Realised spread	**0.111**	**0.166**				
Lambda			**0.005**	**-0.033**	**0.039**	**0.047**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0087**	**0.0081**				
Effective spread	**0.0089**	**0.0081**				
Price impact	**0.0090**	**0.0085**				
Realised spread	**0.0090**	**0.0085**				
Lambda			**0.0000**	**6.0514**	**0.00977**	**0.0125**

Table 5.4: LF vs HF performance – Earnings benchmark period

This table reports the correlations/RMSEs between LF and HF measures across the benchmark period [-60,-11] prior to the earnings announcement. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.817**	**0.412**				
Effective spread	**0.837**	**0.406**				
Price impact	**0.719**	**0.365**				
Realised spread	**0.619**	**0.272**				
Lambda			**0.014**	**-0.073**	**0.268**	**0.227**
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.342**	**0.376**				
Effective spread	**0.359**	**0.375**				
Price impact	**0.272**	**0.280**				
Realised spread	**0.141**	**0.152**				
Lambda			**0.095**	**-0.095**	**0.181**	**0.156**
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.209**	**0.314**				
Effective spread	**0.202**	**0.293**				
Price impact	**0.122**	**0.164**				
Realised spread	**0.105**	**0.166**				
Lambda			**0.007**	**-0.036**	**0.039**	**0.052**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0087**	**0.0082**				
Effective spread	**0.0088**	**0.0082**				
Price impact	**0.0089**	**0.0086**				
Realised spread	**0.0089**	**0.0086**				
Lambda			**0.0000**	**6.1019**	**0.0099**	**0.0126**

Table 5.5: LF vs HF performance – Earnings pre-announcement period

This table reports the correlations/RMSEs between LF and HF measures across the pre-announcement period [-5, 0] prior to/during the earnings announcement. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016. Asterisks indicate significantly different correlations between the benchmark period [-60, -11] and pre-announcement period [-5, 0] at a 5% level of significance.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.576***	**0.475**				
Effective spread	**0.521***	**0.477**				
Price impact	**0.542**	**0.409**				
Realised spread	**0.439***	**0.513**				
Lambda			0.175*	**-0.189***	**0.200**	**0.211**
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.195***	**0.338***				
Effective spread	**0.191***	**0.304***				
Price impact	**0.129***	**0.249***				
Realised spread	**0.111**	**0.125**				
Lambda			0.028	**-0.054**	0.017*	0.029*
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.191***	**0.327***				
Effective spread	**0.189***	**0.306***				
Price impact	**0.113**	**0.196**				
Realised spread	**0.125**	**0.190***				
Lambda			**0.024***	**-0.055**	0.019*	**0.025***
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0096***	**0.0077**				
Effective spread	**0.0097***	**0.0077**				
Price impact	**0.0098***	**0.0079**				
Realised spread	**0.0099***	**0.0080**				
Lambda			**0.0000**	**5.9532**	**0.0092**	**0.0112**

Table 5.6: LF vs HF performance – Earnings post-announcement period

This table reports the correlations/RMSEs between LF and HF measures across the post-announcement period [0, +5] during/following the earnings announcement. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016. Asterisks indicate significantly different correlations between the benchmark period [-60, -11] and post-announcement period [0, +5] at a 5% level of significance.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.627***	**0.356**				
Effective spread	**0.606***	**0.312**				
Price impact	**0.520**	**0.268**				
Realised spread	**0.507***	**0.260**				
Lambda			0.0356	**-0.051**	**0.282**	**0.253**
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.225***	**0.287***				
Effective spread	**0.232***	**0.265***				
Price impact	**0.129***	**0.159***				
Realised spread	**0.153**	**0.161**				
Lambda			0.014	**0.103***	0.051*	0.018*
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.241***	**0.289***				
Effective spread	**0.242***	**0.267***				
Price impact	**0.124**	**0.150***				
Realised spread	**0.151***	**0.151***				
Lambda			0.016	**-0.064**	**0.051***	0.018*
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0089**	**0.0076**				
Effective spread	**0.0088**	**0.0076**				
Price impact	**0.0090**	**0.0079**				
Realised spread	**0.0089**	**0.0079**				
Lambda			**0.0000**	**5.0989**	**0.0089**	**0.0109**

Table 5.7: LF vs HF performance – Earnings event period (positive surprises)

This table reports the correlations/RMSEs between LF and HF measures across the event period [-60,+5] around the earnings announcement for positive earnings surprises. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.581**	**0.432**				
Effective spread	**0.575**	**0.381**				
Price impact	**0.500**	**0.364**				
Realised spread	**0.605**	**0.349**				
Lambda			0.0811	**-0.3011**	**0.32325**	**0.28719**
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.266**	**0.421**				
Effective spread	**0.267**	**0.390**				
Price impact	**0.044**	**0.139**				
Realised spread	**0.196**	**0.216**				
Lambda			**0.06928**	**0.05264**	**0.15132**	**0.16096**
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.178**	**0.311**				
Effective spread	**0.174**	**0.288**				
Price impact	**0.089**	**0.150**				
Realised spread	**0.111**	**0.180**				
Lambda			**0.03068**	**-0.05639**	**0.03366**	**0.0388**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0080**	**0.0074**				
Effective spread	**0.0080**	**0.0073**				
Price impact	**0.0081**	**0.0076**				
Realised spread	**0.0081**	**0.0076**				
Lambda			**0.0002**	**2.7810**	**0.0094**	**0.0123**

Table 5.8: LF vs HF performance – Earnings event period (negative surprises)

This table reports the correlations/RMSEs between LF and HF measures across the event period [-60,+5] around the earnings announcement for negative earnings surprises. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.649**	**0.484**				
Effective spread	**0.661**	**0.454**				
Price impact	**0.559**	**0.409**				
Realised spread	**0.709**	**0.452**				
Lambda			**0.455**	**-0.270**	**0.218**	0.186
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.233**	**0.336**				
Effective spread	**0.254**	**0.334**				
Price impact	**0.120**	**0.183**				
Realised spread	**0.163**	**0.189**				
Lambda			**0.152**	**0.072**	**0.076**	**0.051**
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.191**	**0.327**				
Effective spread	**0.185**	**0.297**				
Price impact	**0.107**	**0.169**				
Realised spread	**0.103**	**0.168**				
Lambda			**0.032**	**-0.078**	**0.045**	**0.056**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0082**	**0.0074**				
Effective spread	**0.0082**	**0.0074**				
Price impact	**0.0083**	**0.0078**				
Realised spread	**0.0084**	**0.0079**				
Lambda			**0.0007**	**2.9357**	**0.0088**	**0.0107**

Table 5.9: LF vs HF performance – Earnings event period (no surprises)

This table reports the correlations/RMSEs between LF and HF measures across the event period [-60,+5] around the earnings announcement for no earnings surprises. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.590**	**0.340**				
Effective spread	**0.599**	**0.336**				
Price impact	**0.584**	**0.307**				
Realised spread	**0.431**	**0.262**				
Lambda			-0.016	-0.138	0.155	0.096
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.360**	**0.148**				
Effective spread	**0.337**	**0.138**				
Price impact	**0.237**	**0.111**				
Realised spread	**0.173**	**0.051**				
Lambda			**0.052**	**-0.117**	**0.100**	**0.125**
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.239**	**0.240**				
Effective spread	**0.229**	**0.220**				
Price impact	**0.153**	**0.139**				
Realised spread	**0.093**	**0.099**				
Lambda			**0.036**	**-0.051**	**0.046**	**0.052**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0108**	**0.0088**				
Effective spread	**0.0109**	**0.0088**				
Price impact	**0.0110**	**0.0093**				
Realised spread	**0.0110**	**0.0092**				
Lambda			**0.0003**	**2.7488**	**0.0092**	**0.0108**

Table 5.10: LF vs HF performance – Earnings event period (large firms)

This table reports the correlations/RMSEs between LF and HF measures across the event period [-60,+5] around the earnings announcement for large firms. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.545**	0.114				
Effective spread	**0.534**	0.054				
Price impact	**0.404**	0.089				
Realised spread	**0.567**	0.013				
Lambda			0.005	-0.112	0.038	0.063
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.240**	**0.367**				
Effective spread	**0.260**	**0.315**				
Price impact	**0.099**	**0.161**				
Realised spread	**0.098**	**0.076**				
Lambda			**-0.103**	0.026	**0.125**	-0.028
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.135**	**0.279**				
Effective spread	**0.130**	**0.234**				
Price impact	**0.055**	**0.105**				
Realised spread	**0.051**	**0.086**				
Lambda			-0.006	**-0.031**	**0.047**	**0.035**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0069**	**0.0053**				
Effective spread	**0.0069**	**0.0053**				
Price impact	**0.0070**	**0.0054**				
Realised spread	**0.0069**	**0.0054**				
Lambda			**0.0001**	**2.4833**	**0.0034**	**0.0035**

Table 5.11: LF vs HF performance – Earnings event period (small firms)

This table reports the correlations/RMSEs between LF and HF measures across the event period [-60,+5] around the earnings announcement for small firms. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.687**	0.174				
Effective spread	**0.673**	0.164				
Price impact	**0.609**	0.198				
Realised spread	**0.635**	0.082				
Lambda			**0.408**	**-0.245**	**0.397**	**0.401**
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.258**	**0.245**				
Effective spread	**0.252**	**0.230**				
Price impact	**0.116**	**0.068**				
Realised spread	**0.166**	**0.198**				
Lambda			**0.201**	**0.072**	**0.180**	**0.219**
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.209**	**0.176**				
Effective spread	**0.203**	**0.154**				
Price impact	**0.119**	**0.083**				
Realised spread	**0.107**	**0.088**				
Lambda			**0.040**	**-0.064**	**0.057**	**0.077**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0109**	**0.0106**				
Effective spread	**0.0109**	**0.0107**				
Price impact	**0.0111**	**0.0111**				
Realised spread	**0.0111**	**0.0111**				
Lambda			**0.0009**	**3.2803**	**0.0144**	**0.0194**

Table 5.12: LF vs HF performance – Earnings event period (pre-GFC)

This table reports the correlations/RMSEs between LF and HF measures across the event period [-60,+5] around the earnings announcement prior to the GFC. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016. The pre-GFC period is divided as the time before 01/2008, the GFC period as 01/2008 until 06/2009 and post-GFC period as the time after 06/2009 (Beber and Pagano, 2013).

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.681**	**0.347**				
Effective spread	**0.666**	**0.312**				
Price impact	**0.579**	**0.300**				
Realised spread	**0.602**	**0.216**				
Lambda			0.079	**-0.369**	**0.425**	**0.428**
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.212**	**0.232**				
Effective spread	**0.219**	**0.214**				
Price impact	**0.164**	**0.152**				
Realised spread	**0.087**	**0.097**				
Lambda			**-0.034**	**-0.144**	**0.049**	**0.040**
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.286**	**0.239**				
Effective spread	**0.283**	**0.224**				
Price impact	**0.195**	**0.193**				
Realised spread	**0.127**	**0.051**				
Lambda			0.0395	**-0.183**	0.097	0.113
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0075**	**0.0067**				
Effective spread	**0.0074**	**0.0067**				
Price impact	**0.0074**	**0.0070**				
Realised spread	**0.0075**	**0.0071**				
Lambda			**0.0001**	**1.9648**	**0.0053**	**0.0065**

Table 5.13: LF vs HF performance – Earnings event period (post-GFC)

This table reports the correlations/RMSEs between LF and HF measures across the event period [-60,+5] around the earnings announcement following the GFC. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016. The pre-GFC period is divided as the time before 01/2008, the GFC period as 01/2008 until 06/2009 and post-GFC period as the time after 06/2009 (Beber and Pagano, 2013).

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.809**	**0.396**				
Effective spread	**0.818**	**0.372**				
Price impact	**0.715**	**0.361**				
Realised spread	**0.702**	**0.252**				
Lambda			**0.112**	-0.377	**0.396**	**0.368**
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.221**	**0.260**				
Effective spread	**0.252**	**0.279**				
Price impact	**0.086**	**0.145**				
Realised spread	**0.202**	**0.174**				
Lambda			-0.129	0.130	0.068	0.077
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.199**	**0.250**				
Effective spread	**0.189**	**0.230**				
Price impact	**0.141**	**0.166**				
Realised spread	**0.090**	**0.117**				
Lambda			0.0327	-0.098	**0.079**	**0.092**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0079**	**0.0073**				
Effective spread	**0.0080**	**0.0074**				
Price impact	**0.0080**	**0.0076**				
Realised spread	**0.0080**	**0.0076**				
Lambda			**0.0002**	**2.8648**	**0.0066**	**0.0089**

Table 5.14: LF vs HF performance – Earnings event period (pre-Chi-X)

This table reports the correlations/RMSEs between LF and HF measures across the event period [-60,+5] around the earnings announcement prior to introduction of Chi-X. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016. The full sample is divided into the period before Chi-X was introduced in Australia (30/10/2011) and the period thereafter.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.665**	**0.606**				
Effective spread	**0.623**	**0.565**				
Price impact	**0.556**	**0.559**				
Realised spread	**0.670**	**0.542**				
Lambda			**0.429**	-0.093	-0.071	-0.137
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.407**	**0.734**				
Effective spread	**0.370**	**0.690**				
Price impact	**0.259**	**0.467**				
Realised spread	**0.207**	**0.407**				
Lambda			**0.157**	-0.031	**0.106**	**0.148**
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.202**	**0.386**				
Effective spread	**0.184**	**0.337**				
Price impact	**0.089**	**0.157**				
Realised spread	**0.108**	**0.204**				
Lambda			**0.017**	**-0.023**	**-0.012**	**-0.014**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0096**	**0.0080**				
Effective spread	**0.0096**	**0.0081**				
Price impact	**0.0097**	**0.0084**				
Realised spread	**0.0097**	**0.0084**				
Lambda			**0.0007**	**3.3082**	**0.0119**	**0.0141**

Table 5.15: LF vs HF performance – Earnings event period (post-Chi-X)

This table reports the correlations/RMSEs between LF and HF measures across the event period [-60,+5] around the earnings announcement following the introduction of Chi-X. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016. The full sample is divided into the period before Chi-X was introduced in Australia (30/10/2011) and the period thereafter.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.810**	**0.390**				
Effective spread	**0.812**	**0.378**				
Price impact	**0.717**	**0.326**				
Realised spread	**0.518**	**0.256**				
Lambda			0.085	**-0.468**	**0.439**	**0.422**
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.510**	**0.239**				
Effective spread	**0.439**	**0.229**				
Price impact	**0.467**	**0.202**				
Realised spread	**-0.289**	**-0.091**				
Lambda			-0.034	**-0.077**	**0.111**	**0.140**
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.226**	**0.257**				
Effective spread	**0.214**	**0.241**				
Price impact	**0.169**	**0.180**				
Realised spread	**0.081**	**0.102**				
Lambda			**0.043**	**-0.183**	**0.092**	**0.111**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0077**	**0.0073**				
Effective spread	**0.0077**	**0.0072**				
Price impact	**0.0077**	**0.0075**				
Realised spread	**0.0078**	**0.0076**				
Lambda			**0.0002**	**2.2904**	**0.0059**	**0.0085**

Table 5.16: LF vs HF performance – Pre-announcement period (positive surprises)

This table reports the correlations/RMSEs between LF and HF measures across the pre-announcement period [-5,0] prior to/during the earnings announcement for positive earnings surprises. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.564**	**0.445**				
Effective spread	**0.569**	**0.391**				
Price impact	**0.532**	**0.307**				
Realised spread	**0.498**	**0.426**				
Lambda			0.071	-0.178	0.179	**0.197**
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.253**	**0.253**				
Effective spread	**0.265**	**0.242**				
Price impact	**0.152**	**0.212**				
Realised spread	**0.177**	**0.078**				
Lambda			0.000	**-0.066**	0.001	0.018
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.174**	**0.313**				
Effective spread	**0.172**	**0.291**				
Price impact	**0.121**	**0.185**				
Realised spread	**0.090**	**0.175**				
Lambda			0.018	**-0.047**	0.011	0.009
Panel D: Average RMSE LF-HF						
Quoted spread	**0.00890**	**0.00691**				
Effective spread	**0.00885**	**0.00683**				
Price impact	**0.00900**	**0.00718**				
Realised spread	**0.00902**	**0.00715**				
Lambda			**0.0001**	**3.2829**	**0.0090**	**0.0114**

Table 5.17: LF vs HF performance – Pre-announcement period (negative surprises)

This table reports the correlations/RMSEs between LF and HF measures across the pre-announcement period [-5,0] prior to/during the earnings announcement for negative earnings surprises. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.253**	**0.481**				
Effective spread	**0.226**	**0.440**				
Price impact	**0.243**	**0.419**				
Realised spread	0.159	**0.369**				
Lambda			**0.210**	-0.117	0.108	0.097
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	0.161	0.445				
Effective spread	0.149	0.395				
Price impact	0.139	0.352				
Realised spread	0.040	0.127				
Lambda			0.032	0.057	0.011	0.025
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.159**	**0.366**				
Effective spread	**0.151**	**0.334**				
Price impact	**0.076**	**0.191**				
Realised spread	**0.098**	**0.194**				
Lambda			0.029	**-0.055**	0.032	**0.043**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0092**	**0.0063**				
Effective spread	**0.0092**	**0.0063**				
Price impact	**0.0091**	**0.0069**				
Realised spread	**0.0092**	**0.0068**				
Lambda			**0.0004**	**3.4482**	**0.0075**	**0.0072**

Table 5.18: LF vs HF performance – Pre-announcement period (no surprises)

This table reports the correlations/RMSEs between LF and HF measures across the pre-announcement period [-5,0] prior to/during the earnings announcement for no earnings surprises. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.445**	**0.361**				
Effective spread	**0.422**	**0.394**				
Price impact	**0.409**	**0.357**				
Realised spread	**0.387**	**0.389**				
Lambda			0.122	-0.170	**0.218**	0.156
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.270**	**0.200**				
Effective spread	**0.276**	**0.186**				
Price impact	**0.128**	**0.183**				
Realised spread	**0.268**	**0.056**				
Lambda			0.048	**-0.103**	0.021	**0.089**
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.229**	**0.235**				
Effective spread	**0.231**	**0.219**				
Price impact	**0.134**	**0.166**				
Realised spread	**0.182**	**0.126**				
Lambda			0.053	**-0.078**	0.021	**0.062**
Panel D: Average RMSE LF-HF						
Quoted spread	0.0106	0.0083				
Effective spread	0.0109	0.0082				
Price impact	0.0109	0.0089				
Realised spread	0.0109	0.0088				
Lambda			0.0002	3.2162	0.0087	0.0100

Table 5.19: LF vs HF performance – Pre-announcement period (large firms)

This table reports the correlations/RMSEs between LF and HF measures across the pre-announcement period [-5,0] prior to/during the earnings announcement for large firms. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.545**	0.211				
Effective spread	**0.539**	0.169				
Price impact	**0.543**	0.244				
Realised spread	**0.454**	0.069				
Lambda			0.029	0.088	-0.191	-0.128
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.149**	**0.367**				
Effective spread	**0.135**	**0.333**				
Price impact	**0.013**	**0.149**				
Realised spread	**0.099**	**0.115**				
Lambda			0.012	-0.021	-0.029	-0.066
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.129**	**0.298**				
Effective spread	**0.121**	**0.263**				
Price impact	**0.073**	**0.162**				
Realised spread	**0.051**	**0.109**				
Lambda			0.014	-0.008	-0.001	-0.030
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0078**	**0.0048**				
Effective spread	**0.0078**	**0.0049**				
Price impact	**0.0080**	**0.0050**				
Realised spread	**0.0080**	**0.0049**				
Lambda			**0.0001**	**2.8890**	**0.0035**	**0.0026**

Table 5.20: LF vs HF performance – Pre-announcement period (small firms)

This table reports the correlations/RMSEs between LF and HF measures across the pre-announcement period [-5,0] prior to/during the earnings announcement for small firms. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.503**	0.125				
Effective spread	**0.428**	0.153				
Price impact	**0.467**	0.113				
Realised spread	**0.299**	0.173				
Lambda			0.193	-0.113	**0.297**	**0.328**
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.258**	**0.223**				
Effective spread	**0.249**	**0.195**				
Price impact	**0.152**	**0.156**				
Realised spread	**0.167**	**0.088**				
Lambda			**0.081**	**-0.075**	0.037	**0.121**
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.207**	**0.201**				
Effective spread	**0.207**	**0.179**				
Price impact	**0.119**	**0.099**				
Realised spread	**0.135**	**0.122**				
Lambda			0.029	**-0.053**	0.024	**0.051**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0120**	**0.0103**				
Effective spread	**0.0120**	**0.0105**				
Price impact	**0.0119**	**0.0110**				
Realised spread	**0.0122**	**0.0109**				
Lambda			**0.0006**	**3.6229**	**0.0138**	**0.0190**

Table 5.21: LF vs HF performance – Pre-announcement period (pre-GFC)

This table reports the correlations/RMSEs between LF and HF measures across the pre-announcement period [-5,0] prior to/during the earnings announcement prior to the GFC. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016. The pre-GFC period is divided as the time before 01/2008, the GFC period as 01/2008 until 06/2009 and post-GFC period as the time after 06/2009 (Beber and Pagano, 2013).

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.535**	**0.398**				
Effective spread	**0.509**	**0.362**				
Price impact	**0.466**	**0.386**				
Realised spread	**0.352**	**0.164**				
Lambda			0.051	**-0.183**	**0.269**	**0.217**
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.088**	**0.063**				
Effective spread	**0.119**	**0.079**				
Price impact	**-0.074**	0.032				
Realised spread	**0.192**	**0.052**				
Lambda			-0.004	0.021	**0.173**	**0.095**
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.226**	**0.287**				
Effective spread	**0.215**	**0.269**				
Price impact	**0.159**	**0.194**				
Realised spread	**0.105**	**0.137**				
Lambda			**0.032**	**-0.101**	**0.082**	**0.091**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0081**	**0.0076**				
Effective spread	**0.0081**	**0.0076**				
Price impact	**0.0082**	**0.0078**				
Realised spread	**0.0081**	**0.0078**				
Lambda			**0.0002**	**2.7358**	**0.0063**	**0.0084**

Table 5.22: LF vs HF performance – Pre-announcement period (post-GFC)

This table reports the correlations/RMSEs between LF and HF measures across the pre-announcement period [-5,0] prior to/during the earnings announcement following the GFC. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016. The pre-GFC period is divided as the time before 01/2008, the GFC period as 01/2008 until 06/2009 and post-GFC period as the time after 06/2009 (Beber and Pagano, 2013).

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.587**	**0.669**				
Effective spread	**0.567**	**0.646**				
Price impact	**0.492**	**0.575**				
Realised spread	**0.501**	**0.553**				
Lambda			-0.109	0.079	-0.096	-0.107
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.247**	**0.437**				
Effective spread	**0.280**	**0.381**				
Price impact	**0.268**	**0.294**				
Realised spread	**0.124**	**0.239**				
Lambda			**-0.139**	0.124	-0.067	-0.022
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.217**	**0.428**				
Effective spread	**0.212**	**0.371**				
Price impact	**0.135**	**0.187**				
Realised spread	**0.118**	**0.258**				
Lambda			-0.029	0.019	-0.050	-0.039
Panel D: Average RMSE LF-HF						
Quoted spread	**0.01208**	**0.00913**				
Effective spread	**0.01208**	**0.00916**				
Price impact	**0.01196**	**0.00934**				
Realised spread	**0.01202**	**0.00929**				
Lambda			**0.0002**	**2.8359**	**0.0139**	**0.0171**

Table 5.23: LF vs HF performance – Pre-announcement period (pre-Chi-X)

This table reports the correlations/RMSEs between LF and HF measures across the pre-announcement period [-5,0] prior to/during the earnings announcement prior to introduction of Chi-X. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016. The full sample is divided into the period before Chi-X was introduced in Australia (30/10/2011) and the period thereafter.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.347**	**0.600**				
Effective spread	**0.308**	**0.583**				
Price impact	**0.250**	**0.513**				
Realised spread	**0.336**	**0.594**				
Lambda			0.193	-0.093	0.055	0.035
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.169**	**0.419**				
Effective spread	**0.174**	**0.382**				
Price impact	**0.090**	**0.290**				
Realised spread	**0.144**	**0.212**				
Lambda			-0.011	0.011	-0.014	-0.013
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.179**	**0.375**				
Effective spread	**0.166**	**0.332**				
Price impact	**0.091**	**0.176**				
Realised spread	**0.108**	**0.222**				
Lambda			0.016	-0.029	-0.023	-0.016
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0100**	**0.0077**				
Effective spread	**0.0101**	**0.0078**				
Price impact	**0.0101**	**0.0081**				
Realised spread	**0.0101**	**0.0080**				
Lambda			**0.0004**	**3.7084**	**0.0110**	**0.0133**

Table 5.24: LF vs HF performance – Pre-announcement period (post-Chi-X)

This table reports the correlations/RMSEs between LF and HF measures across the pre-announcement period [-5,0] prior to/during the earnings announcement following the introduction of Chi-X. See Table 5.2 for variable descriptions. The sample comprises of 222 firms during the period January 1, 2006 to December 31, 2016. The full sample is divided into the period before Chi-X was introduced in Australia (30/10/2011) and the period thereafter.

	HL	CPQS	Amihud	Amivest	HLI	CPQSI
Panel A: Average cross-sectional LF-HF correlations						
Quoted spread	**0.709**	**0.441**				
Effective spread	**0.714**	**0.410**				
Price impact	**0.687**	**0.399**				
Realised spread	**0.719**	**0.405**				
Lambda			0.033	**-0.275**	**0.302**	**0.267**
Panel B: Portfolio time-series LF-HF correlations						
Quoted spread	**0.205**	**0.293**				
Effective spread	**0.174**	**0.266**				
Price impact	**0.157**	**0.232**				
Realised spread	**0.040**	**0.072**				
Lambda			-0.033	**-0.205**	0.009	0.042
Panel C: Individual time-series LF-HF correlations						
Quoted spread	**0.204**	**0.286**				
Effective spread	**0.203**	**0.268**				
Price impact	**0.143**	**0.222**				
Realised spread	**0.128**	**0.127**				
Lambda			**0.037**	**-0.118**	**0.056**	**0.069**
Panel D: Average RMSE LF-HF						
Quoted spread	**0.0091**	**0.0067**				
Effective spread	**0.0090**	**0.0067**				
Price impact	**0.0091**	**0.0070**				
Realised spread	**0.0092**	**0.0071**				
Lambda			**0.0001**	**2.7797**	**0.0062**	**0.0064**

Chapter 6: Conclusion

The process of price discovery in derivative markets and the importance of liquidity are the centrepiece of this work. This research has examined the trading behaviour around security-level and market-wide events, focusing on equity and futures options. Insights have been discovered about the price discovery process of futures and options in commodity markets.

Chapter 2 examines informed options trading around a sample of 352 FDA announcements. This research found that informed traders are aware of the timing of an upcoming announcement at least five days in advance and that at least some informed traders have knowledge of the finer details that affect the price impact of the announcement. Chapter 3 analyses the behaviour of US commodity futures and options IV-based measures as proxies for information leakage around macro-economic and commodity-specific news announcements. This research reports abnormal changes in the levels of futures options IV spreads and skew in the three days preceding news releases. In addition, this research reports a statistically significant relationship between announcement date returns and abnormal changes in pre-announcement IV spreads and skew. Chapter 4 investigates whether commodity futures or options markets are more important in price discovery. New evidence is reported and it is concluded that options lead futures in reflecting new information for most analysed commodities. It was also found that increased speculation, rather than hedging activity, is a key determinant of price discovery in the options markets. Chapter 5 illuminates HF versus LF liquidity measures in times of information asymmetry. This research finds that the closing-percent-quoted-spread is the best proxy for the percent–cost HF benchmarks. In contrast, using cross-sectional, portfolio and individual time-series correlations, the most consistent LF proxies are the HLI and closing-percent-quoted-spread-impact. However, the performance of these proxies weakens in the pre- and post-announcement periods around the earnings announcement.

6.1 Information Leakage and Asymmetry

This research empirically examines intraday and daily options trading data in conjunction with public information, including FDA announcements and macro-economic news to measure abnormal returns. It was found that undisclosed information is incorporated into options data before it is released to the public. Evidence is provided to support the fact that certain market participants trade on non-public information, which is revealed through price movements.

This research also illuminates market liquidity in these times of information asymmetry. The focus is on Australian equities, which highlights a degree of varying liquidity. It is argued that LF measures can be utilised to accurately measure liquidity.

6.2 Price Discovery

This research argues that the recent increase in trading volumes of derivatives leads to a shift in the price discovery process; from traditional instruments such as futures to options. These findings are attributed to the speculative nature of options and ease of trading. However, options markets incorporate a greater amount of noise (i.e., trading frictions). When noise is accounted for, through the ILS estimation methodology, the importance of options in price discovery is emphasised. These findings are supported by prior research conducted by Patel et al. (2018), which finds that the level of noise is crucial to determine the leading market.

6.3 Future Research

Since their inception, derivative markets experienced a significant increase in activity and have become an important area of research. There are increasing liquidity and trading volumes that correspond to an increase in data. This high volume of data affects the feasibility of empirical studies. Increased computing power should be combined with the question of how much data are sufficient and whether simpler statistics are equally representative.

Given the significant advances in trading and technology, future research can examine the prevalence of information leakage in other futures and options markets, such as with interest rates. Further, with the increasing popularity of futures options, increases in electronic trading and changing regulation and market structures, further research could examine the drivers of trading activity and price discovery, especially as other countries and exchanges introduce commodity derivative trading platforms.

With respect to market characteristics, there are several directions for future research, such as liquidity in the presence of insider traders. There is an opportunity to widen the scope of liquidity measures; in particular, adding HF benchmarks. In addition, future research in liquidity and price discovery may consider examining unscheduled news announcements that allow for a cleaner measurement of information leakage and the related information asymmetry.

References

Acharya, V. V. & Johnson, T. C. (2010). More insiders, more insider trading: Evidence from private-equity buyouts. *Journal of Financial Economics, 98*, 500-523.

Acworth, W. (2017). 2017 Annual volume survey - global futures and options volume reaches record level. *Futures Industry Association*.

Adams, Z. & Kartsakli, M. (2017). Have commodities become a financial asset? Evidence from ten years of financialization. *Working paper, University of St Gallen*. *Available at SSRN 2999717.*

Ahern, K. R. (2017). Information networks: Evidence from illegal insider trading tips. *Journal of Financial Economics, 125*, 26–47.

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

Amin, K. I. & Lee, C. M. (1997). Option trading, price discovery, and earnings news dissemination. *Contemporary Accounting Research, 14(2)*, 153-192.

Anthonisz, S. & Putniņš, T. J. (2016). Asset pricing with downside liquidity risks. *Management Science, 63*, 2549-2572.

Augustin, P., Brenner, M., Hu, J. & Subrahmanyam, M. G. (2015). Are corporate spin-offs prone to insider trading? *Available at SSRN 2563012*.

Augustin, P., Brenner, M. & Subrahmanyam, M. G. (2015). Informed options trading prior to M&A announcements: Insider trading? *Available at SSRN 2441606*.

Augustin, P., Brenner, M., Grass, G. & Subrahmanyam, M. G. (2016). How do insiders trade? *Available at SSRN 2848989*.

Back, K., Crotty, K. & Li, T. (2018). Identifying information asymmetry in securities markets. *Review of Financial Studies, 31*(6), 2277–2325.

Baillie, R. T., Booth, G. G., Tse, Y. & Zabotina, T. (2002). Price discovery and common factor models. *Journal of Financial Markets, 5*(3), 309-321.

Bainbridge, S. M. (1998). Insider trading: an overview. *Available at SSRN 132529*.

Ball, R. & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research, 6*, 159-178.

Barber, B., De George E., Lehavy, R. & Trueman, B. (2013). The earnings announcement premium around the globe. *Journal of Financial Economics, 108*, 118-138.

Barraclough, K., Robinson, D. T., Smith, T. & Whaley, R. (2013). Using option prices to infer overpayments and synergies in M&A transactions. *Review of Financial Studies, 26*, 695-722.

Basu, P. & Gavin, W. T. (2010). What explains the growth in commodity derivatives? *Federal Bank of St. Louis Review, 93*(1), 37-48.

Beber, A. & Pagano, M. (2013). Short-selling bans around the world: Evidence from the 2007–09 crisis. *Journal of Finance, 68*, 343-381.

Benet, B. A. & Luft, C. F. (1995). Hedge performance of SPX index options and S&P 500 futures. *Journal of Futures Markets, 15*(6), 691.

Bernile, G., Hu, J. & Tang, Y. (2016). Can information be locked up? Information leakage ahead of macro-news announcements. *Journal of Financial Economics, 121*, 496-520.

Bhattacharya, U. & Daouk, H. (2002). The world price of insider trading. *Journal of Finance, 57*(1), 75-108.

Black, F. (1975). Fact and fantasy in the use of options. *Financial Analysts Journal, 31*(4), 36-41.

Black, F. & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*(3), 637-654.

Bosch, J. & Insup, L. (1994). Wealth effects of Food and Drug Administration (FDA) decisions. *Managerial and Decision Economics, 15*, 589–599.

Boyd, N. & Locke, P. (2014). Price discovery in futures and options markets. *Journal of Futures Markets, 34*, 853–867.

Brennan, M., Huh, S. & Subrahmanyam, A. (2018). High-frequency measures of informed trading around corporate announcements. *Review of Financial Studies, 31*(6), 2326–2376.

Brennan, M., Chordia, T., Subrahmanyam, A. & Tong, Q. (2012). Sell-order liquidity and the cross-section of expected stock returns. *Journal of Financial Economics, 105*, 523-541.

Brogaard, J., Hendershott, T. & Riordan, R. (2014). High-frequency trading and price discovery. *Review of Financial Studies, 27*(8), 2267-2306.

Brunnermeier, M. K. (2005). Information leakage and market efficiency. *Review of Financial Studies, 18*(2), 417-457.

Cao, C., Chen Z. & Griffin, J. M. (2005). Informational content of option volume prior to takeovers. *Journal of Business, 78*, 1073-1109.

Chae, J. (2005). Trading volume, information asymmetry, and timing information. *Journal of Finance, 60*, 413-442.

Chakravarty, S., Gulen, H. & Mayhew, S. (2004). Informed trading in stock and option markets. *Journal of Finance, 59*(3), 1235-1257.

Chan, K., Chung, Y. P. & Fong, W. M. (2002). The informational role of stock and option volume. *Review of Financial Studies, 15*, 1049-1075.

Chan, K., Ge, L. & Lin, T. C. (2015). Informational content of options trading on acquirer announcement return. *Journal of Financial and Quantitative Analysis, 50*, 1057-1082.

Chaney, P. K., Devinney, T. M. & Winer, R. S. (1991). The impact of new product introductions on the market value of firms. *Journal of Business, 573–610*.

Chari, V. V. & Christiano, L. (2017). Financialization in commodity markets. *National Bureau of Economic Research, No.w23766*.

Chen, Y., Gau, Y. & Liao, W. (2016). Trading activities and price discovery in foreign currency futures markets. *Review of Quantitative Finance and Accounting, 46*, 793–818.

Cheng, I. H. & W., X. (2014). Financialization of commodity markets. *Annual Review of Financial Economics, 6*, 419-441.

Christophe, S. E., Ferri, M. G. & Hsieh, J. (2010). Informed trading before analyst downgrades: Evidence from short sellers. *Journal of Financial Economics, 95*, 85-106.

Chung, K. & Zhang, H. (2014). A simple approximation of intraday spreads using daily data. *Journal of Financial Markets, 17*, 94-120.

CME Group Inc., C. G. (2016). Infographic: The remarkable growth of options on futures. Retrieved from http://openmarkets.cmegroup.com/9704/infographic-the-remarkable-growth-of-options-on-futures.

Cohen, K. J., Hawawini, G. A., Maier, S. F., Schwartz, R. A. & Whitcomb, D. K. (1986). Friction in the trading process and the estimation of systematic risk. *Journal of Financial Economics, 12*(2), 263-278.

Collin-Dufresne, P. & Fox, V. (2015). Do prices reveal the presence of informed trading? *Journal of Finance, 70*, 1555-1582.

Corwin, S. & Schultz, P. (2012). A simple way to estimate bid-ask spreads from daily high and low prices. *Journal of Finance, 67*, 719-759.

Cox, J. C. & Ross, S. A. (1979). Option pricing: A simplified approach. *Journal of Financial Economics, 7*(3), 229-263.

Creighton, A., Gower, L. & Richards, A. J. (2007). The impact of rating changes in Australian financial markets. *Pacific-Basin Finance Journal, 15*, 1-17.

Cremers, M. & Weinbaum, D. (2010). Deviations from put-call parity and stock return predictability. *Journal of Financial and Quantitative Analysis, 45*, 335–367.

Dimpfl, T., Flad, M. & Jung, R. C. (2017). Price discovery in agricultural commodity markets in the presence of futures speculation. *Journal of Commodity Markets, 5*, 50–62.

Easley, D. & O'Hara, M. (1987). Price, trade size, and information in securities markets. *Journal of Financial Economics, 19*(1), 69-90.

Easley, D. & O'Hara, M. (1998). Option volume and stock prices: Evidence on where informed traders trade. *Journal of Finance, 53*(2), 431-465.

Elder, J., Miao, H., Ramchander, S. (2012). Impact of macroeconomic news on metal futures. *Journal of Banking & Finance, 36*, 51-65.

Ellis, K., Michaely, R. & O'Hara, M. (2000). The Accuracy of Trade Classification Rules: Evidence from Nasdaq. *Journal of Financial and Quantitative Analysis, 35*(4), 529–551.

Fama, E. F. & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy, 81*, 607-636.

Fernandes, N. & Ferreira, M. A. (2008). Insider trading laws and stock price informativeness. *Review of Financial Studies, 22*(5), 1845-1887.

Figuerola-Ferretti, I. & Gonzalo, J. (2010). Modelling and measuring price discovery in commodity markets. *Journal of Econometrics, 158*(1), 95-107.

Fleming, J., Ostdiek, B & Whaley, R. E. (1996). Trading costs and the relative rates of price discovery in stock, futures, and option markets. *Journal of Futures Markets, 16*, 353–387.

Fleming, M. J. & Remolona, E. M. (1999). Price formation and liquidity in the US Treasury market: The response to public information. *Journal of Finance, 54*, 1901-1915.

Fong, K. Y. L., Holden, C. W. & Trzinka, C. A. (2017). What are the best liquidity proxies for global research? *Review of Finance, 21*(4), 1355–1401.

French, K. R. & Roll, R. (1986). Stock return variances: The arrival of information and the reaction of traders. *Journal of Financial Economics, 17*(1), 5-26.

Garbade, K. D. & Silber, W. L. (1983). Price movements and price discovery in futures and cash markets. *Review of Economics and Statistics, 289–297*.

Ge, L., Lin, T. C. & Pearson, N. D. (2016). Why does the option to stock volume ratio predict stock returns? *Journal of Financial Economics, 120*, 601-622.

Gharghori, P., Maberly, E. D. & Nguyen, A. (2017). Information leakage around stock split announcements: Evidence from the option market. *Journal of Financial and Quantitative Analysis, 52*, 705-735.

Giot, P. (2003). The information content of implied volatility in agricultural commodity markets. *Journal of Futures Markets, 23*, 441-454.

Glosten, L. R. & Milgrom, P. A. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics, 14*, 71-100.

Gonzalo, J. & Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics, 13*, 27–35.

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

Goyenko, R. Y., Holden, C. W. & Trzinka, C. A. (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics, 92*, 153-181.

Griffin, J., Hirschey, N. & Kelly, P. (2011). Why does the reaction to news announcements vary across countries? *Review of Financial Studies, 24*, 3941-3992.

Grossman, S. J. (1986). An analysis of the role of insider trading on futures markets. *Journal of Business, 59*, 129-146.

Grossman, S. J. & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review, 70*(3), 393-408.

Grünbichler, A., Longstaff, F. A. & Schwartz, E. S. (1994). Electronic screen trading and the transmission of information: An empirical examination. *Journal of Financial Intermediation, 3*(2), 166-187.

Hao, Q. (2016). Is there information leakage prior to share repurchase announcements? Evidence form daily options trading. *Journal of Financial Markets, 27*, 79-101.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *Journal of Finance, 50*, 1175–1199.

Hasbrouck, J. (2009). Trading costs and returns for US equities: The evidence from daily data. *Journal of Finance, 64*, 1445-1477.

Hauptfleish, M., Putniņš, T. J. & Lucey, B. (2016). Who sets the price of gold? London or New York. *Journal of Futures Markets, 36*, 564-586.

Hayek, F. A. (1945). The use of knowledge in society. *American Economic Review, 35*(4), 519-530.

Hayunga, D. K. & Lung, P. P. (2014). Trading in the options market around financial analysts' consensus revisions. *Journal of Financial and Quantitative Analysis, 49*, 725–747.

Henry, T. R. & Koski, J. L. (2010). Short selling around seasoned equity offerings. *Review of Financial Studies, 23*, 4389-4418.

Hess, D., Huang, H. & Niessen, A. (2008). How do commodity futures respond to macroeconomic news? *Financial Markets and Portfolio Management, 22*, 127-146.

Hsu, C. H., Chian, Y. C. & Liao, T. L. (2013). Overreaction and underreaction in the commodity futures market. *International Review of Accounting, Banking and Finance, 5*, 61-83.

Hu, J. (2014). Does option trading convey stock price information? *Journal of Financial Economics, 111*, 625-645.

Jayaraman, N., Frye, M. B. & Sabherwal, S. (2001). Informed trading around merger announcements: An empirical test using transaction volume and open interest in options market. *Financial Review, 36*, 45–74.

Jin, W., Livnat, J. & Zhang, Y. (2012). Option prices leading equity prices: Do option traders have an information advantage? *Journal of Accounting Research, 50*, 401-432.

Johnson, T. L. & So, E. C. (2012). The option to stock volume ratio and future returns. *Journal of Financial Economics, 106*, 262–286.

Kacperczyk, M. T. & Pagnotta, E. (2018). Chasing private information. *Available at SSRN 2695197*.

Khan, M. & Lu, H. (2013). Do short sellers front-run insider sales. *Accounting Review, 88*, 1743-1768.

Krinsky, I. & Lee, J. (1996). Earnings announcements and the components of the bid-ask spread. *Journal of Finance, 51*, 1523-1535.

Kusnadi, Y. (2015). Insider trading restrictions and corporate risk-taking. *Pacific-Basin Finance Journal, 35*, 125-142.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Journal of the Econometric Society*, 1315-1335.

Lee, C. M. C. & Ready, M. J. (1991). Inferring trade direction from intraday data. *Journal of Finance, 46*, 733-746.

Lehmann, B. N. (2002). Some desiderata for the measurement of price discovery across markets. *Journal of Financial Markets, 5*(3), 259-276.

Lien, D. & Tse, Y. K. (2000). Hedging downside risk with futures contracts. *Applied Financial Economics, 10*(2), 163-170.

Lin, T. & Lu, X. (2015). Do short-sale costs affect put options trading? Evidence from separating hedging and speculative shorting demands. *Review of Finance, 20*, 1911-1943.

Lo, A. W. & Wang, J. (2006). Trading volume: Implications of an intertemporal capital asset pricing model. *Journal of Finance, 61*, 2805-2840.

Lucia, J. J. & Pardo, A. (2010). On measuring speculative and hedging activities in futures markets from volume and open interest data. *Applied Economics, 42*, 1549–1557.

Luo, X., Qin, S. & Ye, Z. (2016). The information content of implied volatility and jumps in forecasting volatility: Evidence from the Shanghai gold futures market. *Finance Research Letters, 19*, 105-111.

Malkiel, B. G. & Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance, 25*(2), 383-417.

Manaster, S. & Koehler, G. (1982). The calculation of implied variances from the Black-Scholes model: A note. *Journal of Finance, 37*(1), 227-230.

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

Mixon, S. (2009). Option markets and implied volatility: Past versus Present. *Journal of Financial Economics, 94*, 171-191.

Mixon, S. (2011). What does implied volatility skew measure? *Journal of Derivatives, 18*, 9-25.

Muravyev, D., Pearson, N. D. & Broussard, J. P. (2013). Is there price discovery in equity options? *Journal of Financial Economics, 107*, 259–283.

O'Hara, M. (2003). Presidential address: Liquidity and price discovery. *Journal of Finance, 58*(4), 1335-1354.

O'Hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics, 116*(2), 257-270.

Pan, J. & Poteshman, A. M. (2006). The information in option volume for future stock prices. *Review of Financial Studies, 19*, 871-908.

Patel, V., Putniņš, T. J., Michayluk, D. & Foley, S. (2018). Price discovery in stock and options markets. *Available at SSRN 3160578.*

Patel, V. & Michayluk, D. (2016a). Do divestitures create value? Evidence from options markets. *Working paper, University of Technology Sydney.*

Patel, V. & Michayluk, D. (2016b). Return predictability following different drivers of large price changes. *International Review of Financial Analysis, 45*, 202-214.

Peck, A. E. (1980). The role of economic analysis in futures market regulation. *American Journal of Agricultural Economics, 62*, 1037–1043.

Pevzner, M., Xie, F. & Xin, X. (2015). When firms talk, do investors listen? The role of trust in stock market reactions to corporate earnings announcements. *Journal of Financial Economics, 117*, 190-223.

Podolski, E. J., Truong, C. & Veeraraghavan, M. (2013). Informed options trading prior to takeovers: Does the regulatory environment matter? *Journal of International Financial Markets, Institutions and Money, 27*, 286-305.

Putniņš, T. J. (2013). What do price discovery metrics really measure? *Journal of Empirical Finance, 23*, 68–83.

Riordan, R., Storkenmaier, A. & Wagener, M. (2013). Public information arrival: Price discovery and liquidity in electronic limit order markets. *Journal of Banking & Finance, 37*(4), 1148-1159.

Roache, S. K. & Rossi, M. (2010). The effects of economic news on commodity prices. *Quarterly Review of Economics and Finance, 50*, 377-385.

Roll, R., Schwartz, E. & Subrahmanyam, A. (2010). O/S: The relative trading activity in options and stock. *Journal of Financial Economics, 96*, 1-17.

Sanders, D. R., Irwin, S. H. & Merrin, R. P. (2010). The adequacy of speculation in agricultural futures markets: Too much of a good thing? *Applied Economic Perspectives and Policy, 32*, 77–94.

Sammann, D. (2015). Infographic: The remarkable growth of options on futures. *Chicago Mercantile Exchange*.

Sarkar, S. K. & de Jong, P. J. (2006). Market response to FDA announcements. *Quarterly Review of Economics and Finance, 46*, 586–597.

Sarwar, G. (2005). The informational role of option trading volume in equity index options markets. *Review of Quantitative Finance and Accounting, 24*(2), 159-176.

Schestag, R., Schuster, P. & Uhrig-Homburg, M. (2016). Measuring liquidity in bond markets. *Review of Financial Studies, 29*, 1170-1219.

Sharma, A. & Lacey, N. (2004). Linking product development outcomes to market valuation of the firm: The case of the US pharmaceutical industry. *Journal of Product Innovation Management, 21*, 297–308.

Simon, M. (2014). Options on futures: A market rimed for further expansion. *Tabb Group*.

Tang, K. & Xiong, W. (2012). Index investment and the financialization of commodities. *Financial Analysts Journal, 68*(5), 54-74.

Theil, H. (1966). Applied economic forecasts. *Applied Economic Forecasts. North-Holland, Amsterdam.*

Tokic, D. (2013). Crude oil futures markets: Another look at traders' positions. *Journal of Derivatives & Hedge Funds*, *19*(4), 321–342.

Torabzadeh, K. M., Woodruff, C. G. & Sen, N. (1998). FDA decisions on new drug applications and the market value of pharmaceutical firms. *American Business Review, 16*(42).

Wang, C. (2004). The effect of net positions by type of trader on volatility in foreign currency futures markets. *Journal of Futures Markets, 22*, 427-450.

Wolf, A. (1982). Fundamentals of commodity options on futures. *Journal of Futures Markets, 2*, 391-408.

Working, H. (1953). Hedging reconsidered. *Journal of Farm Economics, 35*, 544–561.

Working, H. (1960). Note on the correlation of first differences of averages in a random chain. *Econometrica, 28*(4), 916-918.

Xing, Y., Zhang, X. & Zhao, R. (2010). What does the individual option volatility smirk tell us about future equity returns? *Cambridge University Press, 45*, 641-662.

Yan, B. & Zivot, E. (2010). A structural analysis of price discovery measures. *Journal of Financial Markets, 13*, 1–19.

Yang, J. & Bessler, D. A. (2004). The international price transmission in stock index futures markets. *Economic Inquiry, 42*(3), 370-386.