Essays on Correlated Information Flows and Asset Pricing

Xu Hong - 22996079

BCom (Hons)

Master of Banking and Finance

A thesis submitted in fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Banking and Finance

Monash Business School

Monash University

March 2020

Copyright Notice

Abstract

Recent developments in asset-pricing theories suggest that investors have limited attention and are subject to constraints in processing information. Motivated by the literature, this thesis studies investor attention and information processing, and their asset-pricing implications in different contexts. Chapter 2 examines attention comovement in cross-listed stocks that share the same fundamentals. After establishing the existence of attention comovement in cross-listed stock pairs, the chapter studies the potential determinants of attention comovement. The results suggest that firms' information environment, information shocks and aggregate market attention play a significant role in attention comovement, which provides empirical evidence for both the rational and behavioural views on investor attention. Finally, the chapter shows that correlated attention has an important capital market consequence in that it reduces price deviations in cross-listed stock pairs.

To better understand how information is processed across stocks, Chapter 3 studies news spillover within an industry. Veldkamp (2006) suggests that investors use information from a common subset of assets to value other assets when making investment decisions. Building on Veldkamp's (2006) theoretical prediction and motivated by industry practice whereby investors use bellwether firms to value other firms, the chapter investigates how news is transmitted between industry bellwether firms and peer firms. The results show that bellwether firms' news exhibits significant influence on industry peers' stock prices, trading activity and analyst forecasts. Also, news from bellwether firms contributes more to news of other firms compared to their industry peers. This intra-industry information production affects return comovement. Firms with more informative news are associated with stronger return comovement with the market.

Chapter 4 examines similar research issues in the context of style investing. It investigates whether information at style level contributes to the documented style-related return predictability. Building on Barberis and Shleifer's (2003) style chasing model, the chapter examines how investors allocate attention across different style portfolios. The results show that both style- and firm-level attention is significantly affected by prior style performance. Also, style-level attention contributes to within-style excess return comovement and autocorrelation in style returns.

Overall, this thesis suggests that the arrival of value-relevant information, investors' social interaction and category learning are all plausible explanations of investor attention. Also, it shows that the processing of information has significant asset-pricing implications. Specifically, correlated attention reduces price disparity in cross-listed stocks, intra-industry news spillover affects return comovement, and style-level attention contributes to style-related return predictability.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signature:

Print Name: Xu Hong

Date: 12 December 2019

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisors, Professor Philip Gray and Dr Daniel Chai for their enormous support throughout my PhD course. Your unwavering support, enduring patience, inspiration and encouragement made this thesis possible. I am extremely privileged to have had you both as my supervisors. To Professor Philip Gray, I am deeply indebted to you for many things. Your rigorous scholarship and professionalism in teaching make you the role model of my academic career. Your humbleness, integrity and great sense of humour make you the most approachable and reliable person I always turn to for advice when I experience difficulties in my candidature period. To Dr Daniel Chai, I cannot be more grateful for all the help you have offered me. Thank you for opening my eyes to the world of research, for the inspiration and constant guidance on improving my thesis, and for the continued support and encouragement. To me, you are not only a scholarly mentor but also a good friend.

I would like to express my gratitude to Professor Christine Brown, Professor Stephen Brown, Professor Abe De Jong, Professor Chris Veld, Professor Michael Skully, Professor Yulia Merkoulova, Associate Professor Silvio Contessi, Associate Professor Philip Gharghori, Associate Professor Barry Williams, Dr Thanh Huynh, Dr Hue Hwa Au Yong, and Dr Minh Do for their insightful suggestions to my research.

I would also like to thank my friends Andrew Dallimore, Yanjun Liu, Pham Viet Anh, Tim Kooijmans and Mars Chen. Thank you for the stimulating discussions, for the mutual support, and for all the fun we have had throughout our PhD journey.

Last but not least, I thank my parents and grandparents. Thank you for the unconditional love, support, encouragement and always being there for me. Without you, the work could not have been completed.

Table of Contents

List of Tables

Chapter 1

Introduction

1.1 Background and motivation

The acquisition and dissemination of information are central activities in financial markets. Asset-pricing theories based on efficient markets assume that the diffusion of every type of public information takes place instantaneously among all investors and investors act on the information as soon as it received. In practice, the distillation of new information requires investors' close attention to process all available information and to incorporate it into their investment decisions. Important news or information would not be incorporated into prices unless investors pay attention to it. Merton (1987) highlights incompleteness in dissemination of information and suggests that financial models based on frictionless markets and complete information inadequately capture the complexity of rationality in action.

Attention is a scarce cognitive resource (Kahneman, 1973), as attention to one task necessarily requires a substitution of cognitive resources from other tasks. With respect to investment decisions, given the vast amount of information available, investors must be selective in information processing. Investors' limited attention and computational capacity has motivated a number of behavioural theories that depart from the classical assumptions of strict rationality and unlimited information processing capacity on the part of investors.

Hong and Stein (1999) build a theoretical model that assumes investors are only able to process a small subset of available information and information diffuses gradually across the population of investors. Barberis and Shleifer (2003), Peng and Xiong (2006), Veldkamp (2006) and Mondria (2010) study heuristics that simplify problem solving. Barberis and Shleifer (2003) suggest that, to process vast amounts of information efficiently, investors allocate investment based on exogenous asset styles and simultaneously move in and out of a style depending on its recent performance. Peng and Xiong (2006) show that limited attention leads to category-learning behaviour. An attention-constrained investor tends to allocate more attention to market- and sector-level factors than to firm-specific factors. Veldkamp (2006) and Mondria (2010) argue that, due to information processing constraints, investors choose to observe signals that are good predictors of many assets.

The aforementioned studies recognize investors' limitations in processing information and provide helpful insights to better understand financial behaviour that is otherwise seen as anomalous to standard frictionless-market models. For example, Hong and Stein (1999) show that gradual information diffusion generates short-term underreaction and long-term overreaction in asset returns. Veldkamp (2006) and Mondria (2010) suggest that the use of a common subset of information leads to excess return comovement.

Motivated by the literature, this thesis studies investor attention and information processing, and their asset-pricing implications. The motivations for this thesis are as follows. First, there is growing evidence that investor attention is associated with the pricing of stocks. Existing theories provide two different views on investor attention. The rational view asserts that investor attention is information driven, and is triggered by the arrival of new information (Peng, 2005; Peng and Xiong, 2006; Sims, 2003). The behavioural view contends that attention is partly socially driven, since investors collectively focus on similar firms, and systematically seek out information for similar categorical stocks (Barberis, Shleifer, and Wurgler, 2005; Hirshleifer and Teoh, 2009; Hirshleifer, 2015). As empirical evidence from testing these explanations is limited, the underlying drivers of investor attention are not well understood. This motivates this thesis to examine attention comovement in cross-listed stock pairs that share the same fundamentals and are simultaneously traded in foreign and domestic markets.

The second motivation is to further understand how attention is allocated across stocks. Veldkamp (2006) develops a framework in which investors use information from a subset of assets to value other assets when making investment decisions, and predicts that information production affects return comovement. Specifically, when investors price assets using a common subset of information, news about one asset can affect the price of other assets, generating comovement in asset prices. The source of return comovement has important implications for understanding price formation, asset allocation and risk management. Motivated by Veldkamp's (2006) theoretical framework, this thesis investigates how information is transmitted within an industry, and its implication for return comovement. The finding then motivates an examination of whether information at style level contributes to the documented style-related return predictability.

When making portfolio allocation decisions, many investors categorize assets into different asset classes referred to as styles and move money into and out of these styles. Barberis and Shleifer (2003) suggest that investors' categorical investment and style performance chasing behaviour play an important role in driving style-related return predictability. They argue that investors chase style returns. As a result, higher (lower) returns of a particular style lead to higher fund inflows (outflows). Barberis and Shleifer's (2003) theoretical model predicts excess within-style price comovement and autocorrelation in style-returns. Motivated by their model, the thesis examines how investor attention is affected by past style performance, and whether attention helps explain style-related return patterns.

1.2 Overview of the thesis

This thesis is a comprehensive study that investigates asset-pricing implications of correlated information flows. The thesis has three empirical chapters, presented in Chapters 2 to 4. Each chapter is a stand-alone study that has its own literature review and empirical results, and addresses specific research questions. The chapters are linked and cross-referenced throughout to the purpose of the thesis.

Chapter 2 examines investor attention comovement in cross-listed stocks with the aim of better understanding the driving forces of investor attention. It sheds light on how the processing of information affects price deviations for securities that are fundamentally linked. Chapter 3 investigates intra-industry information transmission and explores an information flow channel for stock return comovement. It demonstrates that intra-industry information production affects return comovement and price efficiency. Chapter 4 further studies attention comovement at style level and shows that collective demand for style-level information contributes to style-related return patterns.

The following subsections provide an overview of each empirical chapter and discuss their academic contributions. Finally, Chapter 5 of this thesis provides a conclusion and discusses directions for future research.

1.2.1 Overview of comovement of investor attention in cross-listed stocks

Using cross-listed stocks as a setting, Chapter 2 examines the extent to which attention on cross-listed stock pairs comoves. To better understand the driving forces of investor attention, we consider both information- and social-related factors in our analysis, and examine how each factor affects attention comovement in cross-listed stock pairs. We then investigate whether attention comovement helps explain deviations from price parity in cross-listed stock pairs. The law of one price states that an identical asset should be traded at the same price regardless of location. Prior studies attribute the existing price deviations to impediments to arbitrage, such as transaction costs, holding costs and capital flow constraints (Gagnon and Karolyi, 2010; Grossmann, Ozuna, and

Simpson, 2007; Suh, 2003). Hong and Stein (1999) suggest that an information imbalance can lead to temporary price divergence of identical securities, when information is impounded into stocks at different speeds. We use attention comovement to capture the relative information diffusion rate in cross-listed pairs, and investigate whether correlated information flows help explain price deviations in cross-listed pairs.

The key findings from Chapter 2 are as follows. First, cross-listed stock pairs exhibit strong attention comovement. The baseline result shows that up to a quarter of the variation in firm attention can be explained by attention on the within-pair counterpart. Second, both information- and social-related factors are important in explaining cross-sectional and time-series variation in attention comovement. The results therefore provide supportive evidence for both rational and behavioural views on investor attention. Third, attention comovement is associated with less deviations from price parity in cross-listed stock pairs.

Chapter 2 makes three main contributions to the literature. First, existing attention literature has predominantly examined the capital market consequences of investor attention without necessarily investigating what drives attention. Our focus on cross-listed stocks allows us to explicitly examine different driving forces of investor attention. Second, the finding that correlated attention is associated with smaller price deviations in cross-listed pairs contributes to the literature on price disparity in the American depositary receipt (ADR) market. Prior studies attempt to explain price deviations using limits to arbitrage and investor sentiment. We show that attention comovement provides incremental explanatory power for price deviations. The finding offers a potential explanation to price deviations between similar assets, such as 'Siamese twin' stocks, closed-end country funds, equity carve-outs and spin-offs. Finally, the study also enriches the newly-established attention comovement literature first documented in Drake, Jennings, Roulstone, and Thornock (2017). Our study extends this strand of literature by applying the attention comovement concept to cross-listed stocks.

1.2.2 Overview of news spillover and return comovement

Chapter 2 suggests that the processing of information has a significant influence on price formation. Chapter 3 extends this line of research by studying intra-industry information transmission and its implication on return comovement. Comovement in asset prices has long been a subject of interest in the literature. However, the source of comovement remains an open question. Theories under the assumptions of no frictions and rational investors suggest that comovement in prices reflects comovement in fundamentals. An alternative view argues that the observed return comovement is too high relative to fundamentals, and favours the friction- and sentiment-based explanations of comovement. Veldkamp (2006) provides a theoretical framework predicting that, in the presence of costly information, investors price assets with a common subset of information. This leads to common movement in asset prices. Motivated by the empirical evidence that investors use information of industry leaders (bellwether firms) to evaluate firms in the same industry, Chapter 3 studies how information is disseminated between bellwether firms and their industry peers, and how intra-industry information production affects return comovement.

The chapter starts with an examination of intra-industry news spillover. If bellwether firms contain information that is useful for other firms in the same industry, news of bellwether firms should be relevant to their industry peers. To test this conjecture, we first investigate whether news on bellwether firms affects industry peers' stock prices, trading activity, and analyst forecasts. We then examine whether bellwether firms' news contributes to news of their industry peers.

Our empirical results show that bellwether firms' news exerts significant influence on their industry peers' stock prices, trading activity and analyst forecast even after controlling for firm-specific news. More importantly, this news spillover is unidirectional. News on non-bellwether firms exhibits no influence on their industry peers. This suggests that news of bellwether firms contains value-relevant information for other firms, above and beyond information from the firms themselves.

Extant literature provides contradictory views on the implication of return comovement. Many studies suggest that low return comovement is an indication of more informative prices, while a growing body of research attributes low return comovement to high information uncertainty. As Dang, Moshirian, and Zhang (2015) note, these contradictory findings are likely driven by the manner in which firm-specific information is measured. The chapter links intra-industry news spillover to this strand of literature. We argue that if a stock's news is important for the pricing of many other stocks, this stock should exhibit stronger return comovement with the market. This is because the information of the stock is capitalized into the prices of many stocks. Consistent with this hypothesis, we find that firms with more contributing news exhibit stronger return comovement. Also, firms with more contributing news are associated with less mispricing. Thus, our findings favor the view that low return comovement implies high information uncertainty.

Overall, Chapter 3 makes two contributions to the literature. First, the finding that investors use news about bellwether stocks to value other stocks validates the empirical prediction of Veldkamp's (2006) theoretical framework. Second, our results add to the long-standing debate on the implication of return comovement by providing direct evidence for a positive relation between return comovement and price informativeness.

1.2.3 Overview of style investing, investor attention and return predictability

Chapter 4 investigates whether information at style level contributes to style-related return predictability. Barberis and Shleifer (2003) predict that style chasing generates excess comovement among assets in the same style, and that style-level momentum and value strategies are more profitable than the asset-level momentum and value strategies. They assume that, to simplify investment decisions, investors allocate funds at style level, and shift funds across the extreme style portfolios based on relative past performance. This implies that styles with extreme performance over the past attract more investor attention. Our empirical analyses support this conjecture. We show that style-level attention is positively related to the absolute prior style returns. Furthermore, a firm's attention is primarily driven by the performance of the style which the firm belongs to rather than the performance of the firm itself.

One of the puzzling empirical findings related to style investment is that assets in the same style comove too much, while assets in different styles comove too little. We link attention comovement to return comovement in order to explore an information flow explanation for this empirical observation. We conjecture that excess return comovement is a result of attention comovement when investors collectively focus on similar categorical stocks. In line with this conjecture, we show that comovement in attention is positively associated with comovement in stock returns. When a stock is reclassified into a new style, both its attention and return comovement with the new (old) style rises (falls). This suggests that return comovement is partially driven by the actions of investors who view individual firms in the context of style categories.

We then investigate whether investor attention contributes to autocorrelation in style returns. Barberis and Shleifer (2003) suggest that style chasing temporarily pushes price away from fundamentals, leading to short-term momentum and subsequent long-term reversals. It is plausible that investors are more likely to allocate funds across styles that attract their attention. Therefore, investor attention should facilitate this observed return pattern. Consistent with this prediction, we document stronger short-term momentum and long-term reversals among more attention-grabbing styles. This suggests that a systematic shift in investor attention contributes to the autocorrelation in style returns.

Chapter 4 sits at the intersection of the investor attention and style investing literatures. It advances the style investing literature from two aspects. First, by showing that prior style performance exerts a significant impact on both style- and firm-level attention, we provide direct evidence for Barberis and Shleifer's (2003) category investment and style chasing prediction. Second, standard asset-pricing models have difficulty explaining style-related return predictability, which calls for the exploration of behavioural explanations. We add to the literature by exploring an information flow channel for the empirical puzzles in style returns. The results suggest that excessive within-style return comovement and autocorrelation in style returns are attributable to the fact that investors process information in the context of style categories. The chapter also contributes to the attention literature by showing that, as a result of category investment, attention is not only a firm-level construct, it is a style construct as well. Also, this macro nature of investor attention has significant capital market consequences.

Chapter 2

Comovement of investor attention in cross-listed stocks

2.1 Introduction

The acquisition of information and its dissemination to other economic units are central activities in capital markets. Standard asset pricing theories based on the efficient market hypothesis assume that new information is immediately diffused across all investors and investors act on the information as soon as it is received (Merton, 1987). However, psychological evidence suggests that the amount of information that can be processed at any time is limited, since attention is a scarce cognitive resource (Kahneman, 1973). Merton (1987) states that investors construct their optimal portfolio by investing in securities of which they are aware. As a result, when there are many alternative options, investors are more likely to consider securities that attract their attention. There has been a growing body of literature investigating the level of investor attention a firm receives and the impact on its stock prices (e.g., Da, Engelberg, and Gao, 2011; Fang and Peress, 2009; Hou, Peng, and Xiong, 2009; Li and Yu, 2012; Lou, 2014).

Recently, Drake et al. (2017) introduce the concept of attention comovement. In particular, Drake et al. (2017) find that the amount of attention a firm receives comoves with the amount of attention paid to its industry and the market as a whole. The study shows that comovement in attention is positively associated with comovement in stock returns. This finding suggests that correlated information flows can lead to comovement in asset prices.

Existing literature provides two possible explanations for attention comovement. The rational explanation asserts that investor attention is information driven, and is triggered by the arrival of new information (e.g., Peng, 2005; Peng and Xiong, 2006; Sims, 2003). Accordingly, assets that experience correlated information shocks exhibit attention comovement. Behavioural theories contend that attention is partly socially driven, since investors collectively focus on similar firms, and systematically seek out information for similar categorical stocks (e.g., Barberis et al., 2005; Hirshleifer and Teoh, 2009; Hirshleifer, 2015). As a result, their attention is affected by the broader level attention paid to the aggregate market. The finding that firm-specific attention comoves with industry and market attention in Drake et al. (2017) is consistent with this socially-driven explanation.

The aim of this chapter is to better understand the driving forces of comovement in investor attention and its implications for asset pricing. For this purpose, we conduct an investigation of cross-listed stocks that have their shares simultaneously traded in the foreign markets and domestic markets.[1] Cross-listed pairs provide an ideal setting to examine the two alternative explanations of attention comovement in that they represent identical claims to the firm's underlying assets, and they have the same exposure to information shocks that are related to firm fundamentals. Thus, correlated information shocks can lead to attention comovement between the cross-listed and home-market shares. Moreover, with increasing interaction of countries through the growth of the international flows of money and investment, attention to cross-listed pairs could be driven by social factors related to the markets where the pairs are listed. Accordingly, there are three research questions to be answered in this chapter. First, to what extent is there attention comovement in cross-listed stock pairs? Second, what are the driving forces of attention comovement? Finally, can attention comovement explain deviations from price parity in cross-listed stocks?

We conduct our analyses using 662 US cross-listed pairs from 36 countries over the 1996-2016 period. Following prior literature, we use trading volume shocks as a proxy for investor attention (Barber and Odean, 2008; Gervais, Kaniel and Mingelgrin, 2001; Hou et al., 2009.[2] Our results provide support for the existence of attention comovement among cross-listed pairs. Time-series regressions show a reliable positive relation between attention in the home market and attention in the US market. We also find an interesting pattern in how attention comovement varies across regions and time periods. Specifically, we find geographic and cultural proximity has a positive effect on attention comovement. Canadian and European firms cross listed in the US market exhibit the strongest attention comovement, while firms with home market domiciled in Asia-Pacific region exhibit the lowest attention comovement. There is an upward trend in attention comovement over our sample period, which we attribute to the improved market integration over time (Brooks and Del Negro, 2004; De Jong and De Roon, 2005). In addition, consistent with Drake et al. (2017), who suggest that aggregate investor attention has a nontrivial impact on variation in firm-specific attention, our result shows a strong positive relation between attention to an individual stock and attention to the aggregate market where the stock is traded.

1 There are a number of reasons why a company chooses to cross-list its shares on one or more foreign stock exchanges. One reason is to increase marketability (Doidge, Karolyi, and Stulz, 2004; Foerster and Karolyi, 1999). Cross-listing enables companies to expand their shareholder base and therefore increases their visibility or recognition in foreign markets.

2 Despite the criticism for its multidimensional characteristics, trading volume is widely used to directly measure the degree of investor attention (Barber and Odean, 2008; Miller, 1977). For cross-listed stocks, trading volume data is readily available compared to other potential proxies for attention.

Having established the existence of attention comovement in cross-listed pairs, we then investigate the determinants of attention comovement. For this purpose, in the spirit of Drake et al. (2017), we regress volume shocks from the home market on volume shocks from the US market. The resulting R^2 captures the extent to which variations in attention can be explained within the pair and is used as a proxy for attention comovement. A higher value of R^2 corresponds to stronger attention comovement.

In our analysis, we consider both information and social-related factors as potential determinants of attention comovement. Specifically, we examine how attention comovement is related to the information environment, information shocks, stock market integration, and aggregate investor attention. In general, our findings provide supportive evidence for both information and socially-driven explanations for attention comovement. At the firm level, attention comovement is positively related to a firm's reporting efficiency, the number of analysts covered, the degree of stock liquidity and the fraction of institutional ownership. At the country level, firms from more developed and liquid home markets tend to exhibit a stronger attention comovement. In addition, firms that experience more frequent information shocks, as proxied by the volatility of stock returns and return on assets (ROA), exhibit stronger attention comovement. The above are consistent with an information based explanation for attention comovement. Also, consistent with a socially-driven explanation, we document a strong positive relation between attention comovement in the home-US pairs and aggregate attention in the US market. This suggests that the aggregate attention spillover from the US market to the home market contributes to the attention comovement in the cross-listed pairs.

Finally, we explore the asset-pricing implications of attention comovement. The law of one price suggests that identical assets should sell for the same price in financial markets due to the workings of arbitrage. The literature shows that market frictions and imperfect information can impede arbitrage, which may result in violations of the law of one price (Lamont and Thaler, 2003; Mitchell, Pulvino, and Stafford, 2002). Previous studies document deviations from price parity for cross-listed pairs (Gagnon and Karolyi, 2010; Grossmann et al., 2007; Suh, 2003). These studies highlight a number of market-friction related factors, such as transaction costs, holding costs and capital flow constraints, as reasons for the mispricing.

The gradual information diffusion hypothesis of Hong and Stein (1999) and subsequent empirical studies suggest that, in the presence of limited information processing capacity, stocks with the same characteristics have a temporary price divergence when information is impounded into stocks at different speeds. Since investor attention is an important factor that results in gradual information diffusion, attention comovement can represent the relative information diffusion rates in cross-listed pairs.[3] Thus, strong (weak) attention comovement should be associated with a relatively small (high) difference in information diffusion rates, resulting in a smaller (larger) deviation from price parity. We test this conjecture in our analysis.

In line with previous studies, our results support the existence of deviations from price parity among cross-listed pairs. On average, we document a daily price deviation of 65 basis points, with the US traded cross-listed shares trading at a premium relative to their home-market counterparts. We investigate the relation between price deviation and attention comovement by performing a panel regression of price deviations on attention comovement. Our results demonstrate a significant negative relation between attention comovement and price deviation. A one-standard deviation increase in attention comovement is associated with a 0.4% decrease in the absolute price deviation, which corresponds to about 17.13% of its standard deviation across all firm-quarters. The negative relation between attention comovement and price deviation remains robust after controlling for factors associated with limits to arbitrage (i.e., holding costs, transaction costs, market volatility, and information asymmetry). Therefore, attention comovement provides incremental predictive power in explaining price deviations in cross-listed stocks.

We further investigate the relation between investor attention from either home or US market and price deviations in cross-listed pairs. The results suggest that unilateral market attention results in large price deviations. This finding is in accord with the view that information imbalance between distinct trading locations can lead to price deviations of identical securities (Chowdhry and Nanda, 1991). More importantly, it further suggests that correlated information flows between different markets are important in driving price parity in cross-listed pairs.

In summary, our study documents the existence of attention comovement in cross-listed stock pairs. We find that information environment, information shocks, and aggregate attention play significant roles in explaining the cross-sectional and time-series variation in attention comovement. This suggests that both information and social factors are important determinants of attention comovement. Finally, we show that attention comovement has important capital market consequences. In particular, we find that correlated attention is associated with smaller deviations from price parity in cross-listed stock pairs.

[3] Hou (2007) empirically shows that slow diffusion of information can result from limited investor attention.

This chapter makes several contributions to the literature. First, it contributes to the emerging literature on investor attention. Existing studies have largely focused on the implications of investor attention for capital markets. However, the underlying drivers of investor attention are not well understood, and empirical evidence is limited. Prior literature provides evidence that investor attention to a specific firm is driven by information shocks that are related to the firm fundamentals (Ben-Rephael, Da, and Israelsen, 2017; Drake, Roulstone, and Thornock, 2012, 2015; Liu and Peng, 2015). The recent work by Drake et al. (2017) argues that, in addition to being an individual construct, attention is also a social construct. The authors show that a substantial portion of firm-level attention is driven by general industry and market attention. Our focus on cross-listed stocks allows us to explicitly examine the different driving forces of investor attention. The empirical evidence supports both the micro- and macro-nature of investor attention.

Second, this chapter contributes to the literature on deviations from price parity in the American depositary receipt (ADR) market. Deviations from price parity in cross-listed stocks are widely documented. Previous studies attempt to explain the price deviations using limits to arbitrage (Eun and Sabherwal, 2003; Gagnon and Karolyi, 2010; Gramming, Melvin, and Schlag, 2005; Grossmann et al., 2007). However, limits to arbitrage cannot explain why price deviations occur in the first place. In this study, we explore an alternative explanation for the existence of price deviations from a channel related to information flows. We show that correlated investor attention provides an incremental explanation for price parity. Our study is also related to several other long-standing empirical puzzles, in which mispricing exists between similar assets. For example, previous studies document price deviations of Siamese twins (Baker, Wurgler, and Yuan, 2012; Froot and Dobora, 1999; Rosenthal and Young, 1990), large closed-end fund discounts or premiums (Bodurtha, Kim, and Lee, 1995; Lee, Shleifer, and Thaler, 1991; Pontiff, 1996), and mispricing in stock carve-outs (Lamont and Thaler, 2003b; Mitchell et al., 2002). The finding of our study sheds light on how the processing of information affects the price deviations for securities that are fundamentally linked.

Finally, the study also enriches the newly-established attention comovement literature. Previous investor attention literature has largely focused on the level of attention that an individual firm receives, and the importance of the level of attention for explaining capital market phenomena. Attention comovement is a new concept introduced by Drake et al. (2017), who argue that attention is not only a firm-level construct but also an industry- and market-level construct. Our study extends this strand of literature by applying the attention comovement concept in cross-listed stocks.

Different from Drake et al.'s (2017) setting that attention to an individual firm comoves with industry and market attention, our study focuses on a single company listed in two different markets with two sets of geographically separate investors that have access to a common set of information. Empirical findings in this study improve our knowledge of capital market implications of attention comovement.

The remainder of this chapter is structured as follows. Section 2.2 reviews the literature on investor attention and cross-listed stocks. Section 2.3 summarises research questions and presents hypotheses. Section 2.4 describes various data sources and methodology. Sections 2.5 and 2.6 present the empirical results and robustness checks, respectively. Section 2.7 concludes the chapter.

2.2 Literature review

2.2.1 The effect of investor attention in asset pricing

Merton (1987) is among the first to demonstrate that investor attention matters for security prices. Merton (1987) refers to the popularity of a stock as its degree of investor recognition and shows that a firm's level of investor recognition is relevant to its cost of capital. In his capital market equilibrium model, Merton assumes that investors know only a subset of the available securities; therefore, less-known firms must offer higher returns to compensate investors for idiosyncratic risk that cannot be diversified away. Consequently, an increase in a stock's investor recognition should lead to a contemporaneous rise in its stock price and lower expected returns in the long run.

In addition to Merton's (1987) investor recognition hypothesis, Barber and Odean's (2008) price pressure hypothesis also predicts that attention affects stock returns. Barber and Odean (2008) propose that attention leads to the net-buying behaviour of individual investors. When individual investors are buying, they face a formidable search problem, since there are a large set of available alternatives. However, when they are selling, this search problem is mitigated since most individual investors hold a relatively small number of stocks in their portfolio, and they do not usually sell short. This means an investor attention shock should lead to, on average, net buying from individual traders. Barber and Odean (2008) predict that the attention-driven buying behaviour results in higher returns over the short run and price reversals over the long run. In brief, Merton's (1987) framework focuses on the asset-pricing implications due to information asymmetry, while Barber and Odean (2008) directly link attention to investors' trading behaviour.

Empirical studies provide supporting evidence for both theoretical predictions. Consistent with the visibility argument in Merton (1987), Gervais et al. (2001) and Kaniel, Ozoguz, and Starks (2012) document a short-term appreciation in stock prices following abnormal trading volume. Fang and Peress (2009) investigate the cross-sectional relation between mass media coverage and expected stock returns, and find that less-covered firms exhibit higher returns after controlling for well-known risk factors. Fang and Peress (2009) attribute the media effect to incomplete diversification as predicted by Merton (1987). Liu, Sherman and Zhang (2014) show that pre-IPO media coverage is negatively related to a stock's expected return for up to three years after its IPO, providing further support for investor recognition hypothesis.

Consistent with Barber and Odean's (2008) price pressure hypothesis, Lou (2014) finds that increased advertising spending is associated with a contemporaneous rise in retail buying and abnormal stock returns, and is followed by lower future returns. Kim and Meschke (2014) show that stocks experience a strong run-up and reversal 11 days after CEO interviews on CNBC. Using internet search volume as a proxy for investor attention, numerous studies document that an increase in Google search volume predicts a significant initial price increase and a subsequent price reversal (Bank, Larch, and Peter, 2011; Da et al., 2011; Joseph, Wintoki, and Zhang, 2011), which provides strong support for the attention-induced price pressure hypothesis.

The primary distinction in empirical findings between Merton (1987) and Barber and Odean (2008) is the return pattern following an attention shock. A longer and more persistent positive return after an attention shock is often attributed to Merton's investor recognition hypothesis, while a short-term price reaction after an attention shock is often attributed to Barber and Odean's (2008) attention theory. The difference in findings might be due to how attention is measured. Most studies testing Merton's model use passive attention measures (for example, abnormal trading volume or media coverage), while most studies testing Barber and Odean (2008) use more active attention measures (for example, online search engine data). Different measures of investor attention will be discussed in Section 2.2.4.

Existing literature also investigates the effect of market-wide attention on stock performance. Vozlyublennaia (2014) measures investor attention using internet search frequency on a broad market index.[4] He finds that attention on an index has a significant short-term effect on the index return. However, different from the prediction by price pressure hypothesis, the author finds that the return could either significantly increase or decrease following a rise in the attention to the index, depending on the nature of the uncovered information. Yuan (2015) proposes Dow record events and front-page market news events as proxies for market-wide attention, and examines the effect of time-varying market attention on the investors' trading behaviour and market returns. Yuan (2015) finds that market-wide attention predicts aggregate net-selling behaviour of individual investors, and consequently, lowers market returns. In a recent work by Andrei and Hasler (2014), the authors assume fluctuations in attention are governed by changes in the state of the economy, and study the role played by investor attention in determining asset prices. Their study theoretically and empirically shows that stock return volatility and the risk premia increase with investor attention. These studies

[4] Vozlyublennaia (2014) argues that internet search on a broad market index better captures attention of retail investors than the search for specific stocks for two reasons. First, the investment choice offered by financial intermediaries typically includes only broad market indexes/portfolios. Second, based on the rational inattention theory (e.g., Peng and Xiong, 2016), retail investors who are more attention constrained are more likely to confine their attention to a broad market index rather than individual stocks.

complement the existing literature that investigates investor attention with a cross-sectional focus.

2.2.2 The effect of investor attention on market efficiency

While much of the capital market theories assume that public information is immediately and costlessly processed by market participants, Grossman and Stiglitz (1980) assert that how informative the price system is depends upon the number of individuals who expend effort to become informed. Therefore, a greater number of informed investors leads to more informative prices, and consequently increases market efficiency. Recent studies provide a theoretical framework in which limited attention can affect price discovery and price informativeness. Sims (2003) applies information theory to study limited attention of an economic agent and its implications for dynamic programming problems in macroeconomics. Hirshleifer and Teoh (2003) and Hirshleifer, Lim, and Teoh (2009) model firms' disclosure policies, and the effects on investor perceptions when investors have limited attention.

A more related study by Peng (2005) examines investor's endogenous attention allocation and provides implications of attention constraints on the cross-sectional differences in the price informativeness. The model predicts that assets with greater total fundamental volatility attract more attention from investors. Since larger stocks tend to attract more investor attention, their prices should incorporate fundamental shocks more quickly and are therefore more informative. This prediction is consistent with empirical findings which document that larger firms' stock prices contain more information about future earnings (Collins, Kothari, and Rayburn, 1987; Freeman, 1987). Peng and Xiong (2006) show that limited attention leads to category-learning behaviour. In particular, an attention-constrained investor tends to allocate more attention to category-level (market and sector) information and less on firm-specific information, and the degree of such categorization depends on the investor's attention constraints. Since stock price becomes more informative when investors process more firm-specific information, Peng and Xiong (2006) illustrate a positive relation between investor attention and price informativeness. Overall, theoretical work suggests that the amount of information incorporated into prices varies across time and across stocks, depending on the investors' attention allocation.

There has been a growing number of empirical studies testing the effect of investor attention on market efficiency. One strand of literature investigates the impact of investor attention on price reaction to news announcements, and documents that market underreaction is associated with proxies for investor inattention. For example, Hirshleifer et al. (2009) find that when there are more firms reporting earnings on the same day, the immediate price reaction to a firm's own earnings surprise is weaker and post-earnings announcement drift is stronger. DellaVigna and Pollet (2009) and Hou et al. (2009) document similar results when announcements are made on Fridays and in down markets. In a cross-sectional test, Hou et al. (2009) show that stocks with higher trading volume experience smaller post-earnings-announcement drift. Similarly, Loh (2010) finds that stocks with higher trading volume react more to stock recommendations during the announcement and experience smaller subsequent price drift. Using internet search volume and information acquisition via EDGAR (Electronic Data Gathering, Analysis, and Retrieval) as proxies for investor attention, Drake et al. (2012, 2015) find that high attention is negatively associated with post-earnings-announcement drift.

The other strand of literature studies the relationship between investor attention and price predictability. Peng and Xiong (2006) show that attention-constrained investors choose to learn first and foremost about the components of returns that are common to multiple stocks. Peress (2014) documents that returns are less dispersed around newspaper strike days. The author argues that the increased return predictability is caused by reduced investor attention since newspaper strikes raise the cost of accessing information. Using Google searches on sports to proxy for exogenous shifts in investor attention, Schmidt (2013) finds that stock prices incorporate less firm-specific news and returns move more synchronously when investor pay more attention to sports. Similar findings are observed by Huang, Huang and Lin (2019) on large jackpot days of Taiwanese nationwide lotteries. In addition, Storms, Kapraun, and Rudolf (2015) show that high retail investor attention leads to better incorporation of idiosyncratic stock information and reduces return predictability, and the effect is more pronounced in bullish markets. At the aggregate level, Vozlyublennaia (2014) finds that attention weakens the predictability of index returns, because more revealed information due to increasing attention improves market efficiency.

Empirical findings provide supporting evidence that investor attention facilitates the price discovery process, reduces price predictability, and improves market efficiency. Some studies further show that the role of attention in market efficiency differs in the cross-section of stocks depending on firm characteristics and investor structure. For example, Hirshleifer et al. (2009) find that distractions affect the market reaction to earnings surprises more strongly for firms with smaller market capitalization, lower analyst following or institutional ownership. Peress (2014) shows that firm news is capitalized more slowly into the returns of small stocks on strike days, while big stocks are insensitive to newspaper strikes. Vozlyublennaia (2014) documents an immediate return response of large stocks to an increase in attention. He attributes the quick response to the availability of ample coverage on the large stocks, which allows investors to process the information quickly. Ben-Rephael et al. (2017) find that institutional attention facilitates the incorporation of information during earnings and recommendation change announcements, while retail attention does not.

2.2.3 Factors associated with investor attention

The previous two sections discuss the role attention plays in financial markets. However, the driving forces for investor attention are not well understood and are only being investigated recently. Drake et al. (2012) investigate the driving forces behind abnormal Google search volume. They find that abnormal search volume is positively associated with corporate news announcements, media articles, and absolute returns, and is negatively associated with investor distraction. Liu and Peng (2015) also use a Google search measure to investigate its pattern and determinants. The study shows that both earnings announcements and macro announcements are associated with greater investor attention, and the increase in attention is highest for the largest firms. However, attention to earnings announcements is lower on days when important macroeconomics news is announced. This suggests that important macroeconomics news shifts investor attention away from analysing firm-specific information, which is consistent with the prediction in Peng and Xiong (2006). Drake et al. (2015) analyse the determinants of information acquisition via EDGAR and find that EDGAR search volume is related to corporate events, previous stock performance, and the strength of the firm's information environment.

Ben-Rephael et al. (2017) explore variables associated with institutional attention and retail attention shocks. They show that firm-specific news is the most important driver of institutional investor attention. In additional, high abnormal stock returns, trading volume, and volatility also trigger high attention from institutional investors. Ben-Rephael et al. (2017) also show that institutional investors pay more attention to larger firms with greater analyst coverage. For retail attention, the results are similar to those obtained from institutional attention. However, the explanatory power of the identified variables is much lower, suggesting that variations in retail attention are more difficult to explain. Overall, consistent with the prediction of the rational inattention model, these empirical studies show that investors actively allocate their attention in response to information shocks, and optimally allocate limited attention across various sources of information (i.e., market-wide vs firm-specific information).

Drake et al. (2017) introduce the concept of attention comovement, which measures the extent to which firm-specific attention is explained by industry- and market-level attention. The study rationalizes attention comovement with two theoretical arguments. First, studies such as Hirshleifer and Teoh (2009) and Hirshleifer (2015) argue that investor learning is at least partially socially driven. As a result of social interaction, investors collectively focus on similar information flows. Therefore, their attention comoves. Second, Barberis et al. (2005) suggest that many investors group assets into categories. Thus, when investors systematically seek out information for similar categorical stocks or experience correlated shocks to the demand for information, investor attention comoves. The empirical results show that up to a quarter of the variation in firm-specific attention can be explained by industry and market attention. Thus, Drake et al. (2017) conclude that investor attention is partly a social process and attention of similar categorical stocks tends to move together. Drake et al. (2017) also investigate asset-pricing implications of attention comovement, and find that attention comovement helps explain the excess comovement in stock returns. They argue that when investor attention comoves with the broader level of attention paid to equity markets, or within industries, the capital allocation may similarly comove, which leads to the observed excess correlations on stock returns.

2.2.4 Proxies for investor attention

Empirical studies on investor attention face the challenge that attention is difficult to observe. Earlier studies use news coverage (Barber and Odean, 2008; Fang and Peress, 2009; Yuan 2015), extreme returns (Barber and Odean, 2008), and abnormal trading volume (Barber and Odean, 2008; Gervais et al., 2001; Hou et al., 2009) as proxies for investor attention. Barber and Odean (2008) argue that attention-grabbing events are likely to be reported in the news, and important news about a firm often results in significant movements in trading volume and returns. Other studies also use price limits (Seasholes and Wu, 2007) and advertising expense (Chemmanur and Yan 2009; Grullon, Kanatas, and Weston 2004; Lou, 2014; Madsen and Niessner, 2016) as a measure of investor attention.

More recent studies propose a number of alternative measures for investor attention. Mondria, Wu, and Zhang (2010) and Da et al. (2011) measure attention allocation using internet search engine data. Da et al. (2011) refer to internet search as a *revealed* attention measure. They argue that if one searches for a stock on the internet, he or she is undoubtedly paying attention to it. Drake et al. (2015) measure attention as the daily number of unique requests for a given firm's filings in the EDGAR (Electronic

Data Gathering, Analysis, and Retrieval). Ben-Rephael et al. (2017) propose a novel measure of institutional investor attention for specific stocks using the news searching and news reading activity on Bloomberg terminals. These measures of investor attention are generally considered advantageous over traditional measures, because they incorporate actively expressed investor interests and more directly capture the actions undertaken by investors to acquire information.

2.2.5 Deviations from price parity for cross-listed stock pairs

Drake et al. (2017) suggest that correlated information flows can lead to comovement in asset prices (discussed in Section 2.2.3). Intuitively, if investor attention from different trading locations to an identical security comoves, it should yield similar price movements. This section discusses the literature on deviations from price parity for cross-listed stock pairs.

The law of one price states that identical goods must have identical prices. In theory, identical securities must have identical prices, and the rule is enforced by arbitrageurs. In practice, however, numerous findings that violate the law of one price have been documented (see, for example, Lamont and Thaler, 2003). Deviations from price parity for cross-listed stock pairs is one example.

Foreign stocks are listed on major US exchanges in several forms: ADRs, New York or Global registered shares (GRS), and direct listings of home-market ordinary shares, such as Canadian cross-listings. ADRs are the most common vehicle through which non-US companies cross list shares in the US. ADRs are US share certificates that represent underlying foreign shares held in custody outside the US. They are traded and settled as conventional shares on the US exchanges, and represent identical claims to the firm's cash flow as their home-market counterparts. From this perspective, the two are arguably the same security (Pulatkonak and Sofianos, 1999). Furthermore, since ADRs are two way convertible into their underlying securities, the law of one price suggests that the pair should trade at parity. However, empirical studies show that deviations from price parity for cross-listed pairs do exist. Existing studies highlight a number of factors which can lead to the price disparity.

Grossmann et al. (2007) examine the determinants of discounts and premiums on the prices of 74 ADRs from nine countries, and show that transaction costs, holding costs, as well as the consumer sentiment in the US contribute to the mispricing. Suh (2003) studies the price movement of ADRs whose home-markets are in emerging countries, and finds that divergence from parity values is observed for the majority of ADRs in the sample. In line with Grossmann et al. (2007), Suh (2003) shows that the

ADRs premium (or discount) is associated with US market sentiment. Rabinovitch, Silva, and Susmel (2003) investigate ADR returns relative to the returns of their respective locally traded shares in Chile and Argentina. The results show that the mispricing of ADRs is negatively related to the level of home market liquidity. Furthermore, consistent with the frequently cited case study of India's Infosys discussed in Puthenpurackal (2006), Rabinovitch et al. (2003) show that capital flow restrictions augment ADR mispricing. Unlike those studies which limit their analysis to a small country sample, more recent work by Gagnon and Karolyi (2010) examines the deviations from price parity for cross-listed pairs based on a sample of 506 US cross-listed stocks from 35 countries. Using daily price data, they document an economically small but highly volatile price deviation between the pairs. Their empirical finding suggests that the magnitude of the price deviation is significantly related to idiosyncratic risk and the information asymmetry. Overall, the result is supportive of costly arbitrage hypothesis.

2.2.6 Limited attention and return predictability in economically linked stocks

Consistent with the view that investors have limited attention, Hong and Stein (1999) assume that, as a result of limited information processing capacity, each type of agent is only able to process a subset of available public information, and many of them may not be able to extract information from asset prices. The gradual diffusion of information suggests that value-relevant information diffuses gradually across informationally segmented markets. Limited investor attention, when combined with gradual diffusion of information, predicts that valuable information that originates from one market reaches investors in the other market only with a lag, generating return predictability. Empirical studies on the return predictability among firms which are economically linked provide supporting evidence for this argument. Specifically, limited investor attention and slow information diffusion contribute to lead-lag effects related to industry factors (e.g., Hong, Torous, and Valkanov, 2007; Hou, 2007) and lead to significantly predictable returns along the supply chain (Cohen and Frazzini, 2008; Menzly and Ozbas, 2010). Recent studies also document the existence of predictable patterns in returns using information about a firm's alliance partners (Cao, Chordia, and Lin, 2016) and its foreign operations (Huang, 2015).

There are a number of studies examining the effect of investor inattention and gradual information diffusion in the context of relative valuation strategy such as pairs trading. The idea behind the pairs trading strategy is to first identify a pair of stocks that show similar historical price movement, and then simultaneously take a long-short position whenever there is sufficient divergence between the prices in the pair, betting the price divergence is temporary and will converge over time. [5] Engelberg, Gao, and Jagannathan (2009) show that when common information diffuses into stocks at different rates, the prices of the two stocks in the pair can temporarily move apart. Consistent with the delay in information diffusion explanation, Chen, Chen and Li (2012) and Jacobs and Weber (2015) document that a pairs trading strategy is more profitable among firms with noisier information environments. They are generally small in size, without media coverage, and have lower investor recognition and lower analyst coverage. In time-series analysis, Jacobs and Weber (2015) show that the success of pairs trading is positively related to investor inattention. In particular, returns from pairs trading are on average 38 bps larger on days with many newly opened pairs as opposed to days with a few newly opened pairs. The return is even larger in down markets as opposed to up markets.[6] Overall, the existing literature provide evidence that limited attention and slow diffusion of information play a significant role in determining the price movement of the economically correlated stocks.

[5] Several strategies are similar to pairs trading. For example, these strategies that consider relative pricing of shares due to differences in trading locations (Froot and Dabora, 1999; Scruggs, 2007) or differences in cashflow rights and voting rights (Smith and Amoako-Adu, 1995; Zingales, 1995; Schultz and Shive, 2008).
[6] Karlsson, Loewenstein, and Seppi (2009) document "ostrich effect" in financial market, which suggests investors pay more attention to stocks in rising markets, but "put their heads in the sand" in flat or falling markets.

2.3 Research questions and statement of hypotheses

2.3.1 Research questions

Motivated by the existing literature, this chapter empirically examines the attention comovement in cross-listed stock pairs. The study aims to investigate the following research questions: (1) Is there attention comovement in cross-listed stock pairs? (2) What determines the cross-sectional and time-series variation in attention comovement? (3) To what extent can attention comovement explain the deviations from price parity for cross-listed stock pairs?

2.3.2 Hypotheses

Prior literature provides evidence that investors actively increase their attention in response to information shocks (e.g., Hirshleifer and Teoh, 2003; Hirshleifer et al., 2009; Peng, 2005; Peng and Xiong, 2006; Sims, 2003). Cross-listed stock pairs represent identical claims to the underlying firm's assets, and they have the same exposure to information shocks that are related to firm fundamentals. Thus, investor attention to the cross-listed pairs should be correlated. Accordingly, our first hypothesis, expressed in the alternative form, is as follows:

H1: Cross-listed stock pairs exhibit attention comovement.

Certain firm and market characteristics are considered potential determinants of attention comovement in cross-listed stocks. Rational inattention literature suggests that investor attention is driven by the arrival of new information. As such, attention to cross-listed pairs should move in the same direction when investors from both home and US markets actively allocate their attention in response to information shocks to the stock. For this reason, we expect that attention comovement is affected by information related factors, such as information environment and the frequency of information shocks.

Behavioural theories argue that, as a result of investors' social interaction, firm-specific attention is largely affected by the broader level of attention paid to the aggregate markets. Therefore, we expect aggregate attention also plays a role in driving attention comovement in cross-listed pairs. To be more specific, we argue that aggregate attention shocks can spill over from US market to home market (or from home market to US market), leading to attention comovement in cross-listed pairs. Hence, we expect larger aggregate attention shocks and more integrated stock markets are associated with stronger attention comovement. Accordingly, our second hypotheses are:

H2a: Attention comovement in cross-listed pairs is positively associated with the degree of information environment and information shocks.

H2b: Attention comovement in cross-listed pairs is positively associated with the degree of aggregate attention and stock market integration.

The existence of deviations from price parity in cross-listed pairs is widely documented. Previous studies suggest that market-friction related factors can explain the price disparity. However, the impediments to price parity due to information flows have not been well explored. If the paired stocks experience different rates of information flows, it is reasonable to expect temporary disparity in their prices. Attention comovement directly captures correlated information flows for cross-listed pairs. Strong attention comovement suggests that information is rapidly diffused to the cross-listed pair at a similar rate, which consequently reduces the likelihood of price deviations. Accordingly, our third hypothesis is:

H3: There is a negative relation between attention comovement and price deviations in cross-listed stock pairs.

2.4 Data and methodology

2.4.1 Data

Our sample begins with a complete list of foreign stocks listed in the US in the form of American Depositary Receipts (ADRs) or in the form of ordinary shares, which are available in the Thomson Reuters Datastream database as of the end of December 2016.[7] Firms that cross list via an ADR have four options to choose from: Level I, Level II, Level III, and Rule 144A programs. The four types of ADRs differ according to their ability to raise capital from US markets and their degree of compliance with governance and disclosure requirements. Level III and Rule 144A offer access to US primary capital markets (i.e., raising capital), whereas Level I and Level II allow access to US secondary markets only. In addition, Levels I, II, and III allow foreign firms to target both public and private US investors, while Rule 144A gives access to US private institutional investors only. Level I ADRs are traded over the counter (OTC), while Level II and Level III ADRs can be traded on NYSE, NASDAQ, or AMEX. Rule 144A ADRs are traded through Automated Linkages (PORTAL) among Qualified Institutional Buyers (QIBs). Since our focus is on exchange-listed ordinary shares, we only include ADRs that are classified as Level II and Level III.[8] We also exclude preferred shares, warrants, any issues denoted as 'Units' or 'Funds', and stocks with no home-market counterparts available in Datastream. After the screening process, our initial sample includes 816 US cross-listed foreign companies.

Data for our empirical analysis are obtained from six different databases. We collect daily closing price, trading volume, market capitalization and market index price for each cross-listed pair from Datastream. Earnings announcements and analyst estimates data are collected from I/B/E/S. Dividend announcements are obtained from CRSP. Accounting data are collected from Compustat. Institutional ownership data are extracted from Thomson Reuters 13F database. We obtain intraday bid and ask quotes of the US-listed shares from Thomson Reuters Tick History (TRTH). Since TRTH has data coverage available from January 1996, our sample therefore starts from January 1996. To ascertain the accuracy of our home-US stock matching process, we cross-reference each ADR issue with the Bank of New York (BONY) and J.P. Morgan's DR directory.[9]

[7] For example, Canadian cross-listings are in the form of ordinary shares.

[8] We do not distinguish the two types of ADRs in the reported results. In the untabulated analysis, the results for Level II and Level III are qualitatively similar. However, compared to Level II ADRs, Level III ADRs exhibit stronger attention comovement. Also, the negative association between attention comovement and price deviation is stronger for Level III ADRs than Level II ADRs.

[9] The data set is available on www.adrbnymellon.com and www.adr.com

We also use these two data sources to validate the ADR bundling ratio (number of home-market shares per ADR) extracted from the Datastream series.

Our key variable of interest, investor attention, is measured by volume shocks. In the spirit of Llorente, Michaely, Saar and Wang (2002), volume shock (V) is computed at the daily level:

$$V_{i,t} = \log(Turnover_{i,t}) - \frac{1}{200}\sum_{s=-200}^{-1} \log\,(Turnover_{t+s}) \tag{2.1}$$

where

$$\log(Turnover_{i,t}) = \log(turnover_{i,t} + 0.00000255)$$

turnover for stock i at day t is defined as the total number of shares traded that day divided by the total number of shares outstanding. A small constant (0.00000255) is added to turnover before taking logarithm to avoid the problem of zero daily trading volume. Volume shock is computed by subtracting a 200-day moving average from the daily log-turnover. As our trading volume data starts from 1st January 1996, the volume shock is available from 8th October 1996.

Since the purpose of this study is to investigate attention comovement in home-US stock pairs, we require that each stock has valid volume shock data in both home and US markets. Furthermore, to ensure sufficient numbers of observations in subsequent analyses, we exclude all firm-quarters with less than 24 valid daily observations. We define a valid observation as one for which the volume shock ($V_{i,t}$) is available in both US and home markets. In the end, our sample consists of 662 firms and 26,820 firm-quarters from 36 countries.

Table 2.1 provides a breakdown of our sample by home country, year, and industry. Of the 662 firms included in our sample, 102 (15.41%) are domiciled in 13 emerging market countries, as defined by International Monetary Fund (IMF). Canada has the largest number of issues with 305 common shares listed in the US market, which accounts for 46.07% of the sample. UK (65) and Israel (43) have the second and third largest issues, respectively, followed by Brazil (26), Mexico (19), Hong Kong (18) and Chile (16).

Panel B summarizes the number of US cross-listings in our sample as well as the annual growth rate from 1996 to 2016. Our sample begins with 170 firms in 1996 and expands to 401 firms in 2016. In general, US cross-listings have experienced a significant increase in late 1990s and early 2000s. However, a noticeable shift occurred in 2004, since then the number of new listings has decreased and witnessed a further slowdown during the financial crisis period. Recent studies attribute the drop in the new US cross-listings to the passage of the Sarbanes-Oxley Act (SOX) in 2002, which makes it more costly for foreign firms to have a US listing (Bianconi, Chen, and Yoshino, 2013; Doidge, Karolyi, and Stulz, 2010; Marosi and Massoud, 2008). Panel C reveals the diverse industrial makeup of our sample, with representation from 53 industry groups based on the two-digit Standard Industrial Classification (SIC) Codes. Metal and mining (109), Chemicals (63), Communications (55), Oil and gas extractions (49), Business services (42), and Depository institutions (41) are among the largest groups represented in our sample.

Table 2.1 Distribution of US cross-listings by home country, year and industry

This table reports the composition of our sample by country of origin, year and industry classification. Our sample is drawn from a complete list of US cross-listed stocks available in Datastream as of the end of 2016 and consists of 662 firms from 36 countries whose shares are listed concurrently at home and in the US on the AMEX, NYSE, or NASDAQ in the form of ADRs or ordinary shares. The industry classification is based on the two-digit Standard Industrial Classification (SIC) Codes. Our sampling period starts on 1st January 1996 and ends on 31st December 2016.

Panel A: By country			Panel B: By year		
Country	Number	Percentage	Year	Number	Growth
Argentina	14	2.11%	1996	170	
Australia	15	2.27%	1997	214	25.88%
Belgium	5	0.76%	1998	247	15.42%
Brazil	26	3.93%	1999	261	5.67%
Canada	305	46.07%	2000	261	0.00%
Chile	16	2.42%	2001	290	11.11%
Colombia	3	0.45%	2002	312	7.59%
Czech Republic	1	0.15%	2003	347	11.22%
Denmark	1	0.15%	2004	366	5.48%
Finland	2	0.30%	2005	377	3.01%
France	12	1.81%	2006	383	1.59%
Germany	7	1.06%	2007	374	-2.35%
Hong Kong	18	2.72%	2008	375	0.27%
India	8	1.21%	2009	378	0.80%
Indonesia	1	0.15%	2010	381	0.79%
Ireland	4	0.60%	2011	388	1.84%
Israel	43	6.50%	2012	388	0.00%
Italy	10	1.51%	2013	392	1.03%
Japan	15	2.27%	2014	401	2.30%
Mexico	19	2.87%	2015	402	0.25%
Netherlands	11	1.66%	2016	401	-0.25%
New Zealand	2	0.30%			
Norway	6	0.91%			
Peru	3	0.45%			
Philippines	1	0.15%			
Russia	2	0.30%			
Singapore	1	0.15%			
South Africa	7	1.06%			
South Korea	8	1.21%			
Spain	7	1.06%			
Sweden	8	1.21%			
Switzerland	8	1.21%			
Taiwan	6	0.91%			
Turkey	1	0.15%			
United Kingdom	65	9.82%			
Venezuela	1	0.15%			
Total	662	100.00%			

Panel C: By industry			
Industry	Number	Industry	Number
Agricultural production	4	Lumber and wood products	1
Agricultural services	1	Metal and mining	109
Amusement and recreation services	1	Miscellaneous retail	1
Building materials and gardening supplies	1	Motion pictures	2
Business services	42	Non-depository institutions	3
Chemical and allied products	63	Non-metallic minerals, except fuels	5
Coal mining	1	Oil and gas extraction	49
Communications	55	Paper and allied products	2
Depository institutions	41	Petroleum and coal products	11
Eating and drinking Places	3	Pipelines, except natural gas	1
Electric, gas, & sanitary services	27	Primary metal industries	15
Electronic and electric Equipment	39	Printing and publishing	7
Engineering and management services	27	Railroad transportation	3
Fabricated metal products	3	Real estate	9
Food and kindred products	13	Rubber and plastics products	3
Food stores	4	Security and commodity brokers	4
Furniture and fixtures	1	Services, not elsewhere classified	2
General building contractors	3	Stone, clay, and glass products	6
Health services	4	Textile mill products	1
Heavy construction, except building	1	Tobacco products	3
Holding and other investment offices	3	Transportation equipment	14
Hotels and other lodging places	3	Transportation services	2
Industrial machinery and equipment	12	Transportation by air	11
Instruments and related products	20	Trucking and warehousing	1
Insurance agents, brokers, and service	2	Water transportation	5
Insurance carriers	8	Wholesale trade	9
Local and interurban passenger transit	1		

2.4.2 Methodology

2.4.2.1 Attention comovement in cross-listed stock pairs

To examine whether attention comovement exists in cross-listed pairs, we perform the following time-series regressions for each of the 26,820 firm-quarters in our sample using all daily observations for each quarter:

$$V_{i,t}^{H} = a_i + \beta_{1i}^{US}V_{i,t}^{US} + \beta_{2i}V_{i,t-1}^{H} + \beta_{3i}V_{i,t-1}^{US} + \beta_{4i,H}V_{H,t} + \beta_{5i,US}V_{US,t} + \beta_{6i,H}V_{H,t-1} + \beta_{7i,US}V_{US,t-1} + \varepsilon_{i,t} \quad (2.2a)$$

$$V_{i,t}^{US} = a_i + \beta_{1i}^{H}V_{i,t}^{H} + \beta_{2i}V_{i,t-1}^{US} + \beta_{3i}V_{i,t-1}^{H} + \beta_{4i,US}V_{US,t} + \beta_{5i,H}V_{H,t} + \beta_{6i,US}V_{US,t-1} + \beta_{7i,H}V_{H,t-1} + \varepsilon_{i,t} \quad (2.2b)$$

where $V_{i,t}$ denotes the volume shock to firm *i* on day *t*. The superscripts '*H*' and '*US*' for the volume shock ($V_{i,t}$) denote those associated with the home-market shares ($V_{i,t}^{H}$) , and those of the US-traded cross-listed shares ($V_{i,t}^{US}$), respectively. $V_{i,t}$ is derived from Equation (2.1).

β_1^{US} from (2.2a) and β_1^{H} from (2.2b) capture the attention comovement of home-US stock pair. In addition, we include the lagged volume shock for both home- and US-market shares ($V_{i,t-1}^{H}$ and $V_{i,t-1}^{US}$) to control for time-series autocorrelations. Since attention comovement in cross-listed pair can be attributed to aggregate attention shocks at home and/or in the US market, we control for these market-wide sources of attention comovement by including contemporaneous and lagged home market volume shock and US market volume shock on the right-hand side of Equations (2.2a) and (2.2b). Home market attention shock (V_H) is computed as equal-weighted volume shocks for all listed stocks within the home market excluding firm *i*. US market attention shock (V_{US}) is computed the same way for all available stocks, excluding stock *i*, from NYSE, Nasdaq and AMEX. Hypothesis 1 predicts there exists attention comovement in cross-listed stock pairs. Therefore, we expect β_1^{US} and β_1^{H} to be significantly positive across firms and quarters.

2.4.2.2 The determinants of attention comovement

The previous section aims to first establish the existence of attention comovement in home-US stock pairs. β_1^{US} and β_1^{H} from the time-series regressions reveal the statistical and economic significance of attention comovement. We now turn to examining the determinants of time-series and cross-sectional variations in attention comovement. To be in line with the literature and for consistency in subsequent analyses, we employ a two-stage methodology analogous to Drake et al. (2017) to measure attention comovement for each pair. In stage one, we estimate the following time-series regression using daily data on a quarterly basis to obtain the adjusted R^2:[10]

$$V_{i,t}^{H} = a_i + \beta_i^{US} V_{i,t}^{US} + \varepsilon_{i,t} \tag{2.3}$$

In stage two, we take the natural logarithm transformation of the adjusted R^2, and attention comovement for firm i over quarter q is defined as:

$$AttentionComove_{i,q} = \ln\left(\frac{R^2}{1-R^2}\right) \tag{2.4}$$

$AttentionComove$ captures the degree of attention comovement in the cross-listed pair. A large value of *AttentionComove* represents a high level of investor attention comovement in the cross-listed pair.

Our analysis examines a set of market and firm characteristics that are potential determinants of attention comovement in cross-listed pairs. Motivated by the information-driven explanation for attention comovement, we consider factors related to a firm's information environment, information shocks, and home-country information environment. Motivated by the socially-driven explanation for attention comovement, we also include aggregate investor attention and market integration. Accordingly, we estimate the following regression across firms and over quarters:

$$\begin{aligned} AttentionComove_{i,q} = a &+ \beta_1 Firm\ information\ environment_{i,q} + \\ &\beta_2 Information\ shocks_{i,q} + \beta_3 Market\ attention_{i,q} + \\ &\beta_4 Market\ integration_{i,q} + \\ &\beta_5 Home\ country\ information\ environment_{i,q} + \\ &\textstyle\sum_{k=6}^{K} \beta_k Control_{i,k,q} + \varepsilon_{i,q} \end{aligned} \tag{2.5}$$

[10] Same R^2 is obtained from $V_{i,t}^{US} = a_i + \beta_i^{H} V_{i,t}^{H} + \varepsilon_{i,t}$. Since we measure the trading volume shock in home-market ($V_{i,t}^{H}$) and US-market ($V_{i,t}^{US}$) on the same calendar date, our measure is subject to time-difference limitation, especially for those stocks domiciled in Asia-pacific markets which have non-overlapping trading hours with the US market. We address this issue in Section 2.6.3.1.

The independent variables are categorised into firm information environment, information shocks, market attention, market integration, home country information environment, and firm characteristics as control variables.

Firm's information environment

The literature argues that information acquisition is a trade-off between the costs of gathering information and the expected benefits of trading on such information (Grossman and Stiglitz, 1980; Diamond and Verrecchia, 1981). The costs and benefits are related to the strength of the information environment of a firm. Since a strong information environment should lead to more active response to information shocks, we expect cross-listed stocks with strong information environment to exhibit a high level of attention comovement. Building on the existing literature, we use earnings announcement speed, number of analysts, institutional ownership, and stock illiquidity as proxies for firm-level information environment. See Appendix A2.2 for details of these specific variables.

Information shocks

Empirical findings suggest that investor attention increases with information shocks. Firms whose returns and fundamentals are more volatile should be associated with a higher frequency of information shocks and hence draw more investor attention. Ben-Rephael et al. (2017) document that institutional investors' attention increases with stock return volatility. Drake et al. (2017) show that firms with a large standard deviation in return on assets (ROA) receive more firm-specific investor attention. As such, we conjecture that cross-listed stocks with high return and ROA volatility are associated with stronger attention comovement.

Home-country information environment

We consider legal environment, financial market development, and market liquidity as proxies for a home market's information environment. We use an anti-director rights index and disclosure requirement index to capture a home country's legal protection, and use the stock market size and turnover to proxy for financial market development and liquidity. These variables are widely used as proxies for information environment (See Appendix A2.2 for details of these specific variables). We expect attention comovement to be positively related to the home market's legal environment, market development, and market liquidity.

Aggregate investor attention

Theoretical and empirical studies provide evidence that limited investor attention leads to category-learning behaviour (Peng and Xiong, 2006; Drake et al., 2017). As a consequence, firm-specific attention is largely affected by the market-wide investor attention. To investigate the extent to which attention comovement in cross-listed pairs is driven by market-wide investor attention, we include the aggregate attention shocks in the US and home markets in our analysis. Aggregate attention shocks for each quarter are calculated by averaging the daily aggregate volume shocks (as defined in Section 2.4.2.1) over the number of days for which data are available during that quarter. We predict that attention shocks in one market can cause the market-wide attention spillover, which leads to stronger firm-level attention comovement. That is, there is a positive relation between attention comovement and aggregate attention shocks.

Market integration

Increasing market integration leads to information transmission across markets (Eun and Shim, 1989; In, Kim, Yoon, and Viney, 2001; Singh, Kumar, and Pandey, 2010). A cross-listed stock whose home market is more integrated with the US market is expected to exhibit stronger attention comovement than a cross-listed stock whose home market is less integrated with the US market. Following prior literature (e.g., Lee and Kim, 1993; Goetzmann, Li, and Rouwenhorst, 2005), we use stock market return correlations to capture market integration. We expect a positive relation between market correlation and attention comovement in cross-listed pairs.

Control variables

We control for firm size, price, book-to-market and lagged attention comovement in our regression analysis. Size is the natural logarithm of market capitalization expressed in US dollars at the beginning of each quarter. Price is the natural logarithm of home-market share price expressed in US dollars at the beginning of each quarter. Book-to-market is defined as the balance sheet value of the common equity divided by the market value of the common equity extracted from Datastream on a quarterly basis and is natural logged. Lagged attention comovement is the attention comovement over the previous quarter.

2.4.2.3 The effect of attention comovement on price deviations

We follow Gagnon and Karolyi (2010) to measure deviations from price parity in cross-listed pairs after adjusting for currency and bundling ratio at a daily basis. One empirical challenge in comparing prices of cross-listed pair is that the trading hours for the majority of home markets are different from the US market.[11] For example, many Latin American stock markets are imperfectly synchronized to the US trading hours, being off by 1-4 hours. European, African, and Middle Eastern stock markets trade with 5 or 6 hours ahead of the US. Asia-pacific markets trade at least 12 hours ahead and thus, have no overlapping trading hours with the US markets.

In order to overcome the problem of non-synchronous trading hours and time zone differences, we follow Gagnon and Karolyi (2010) to match US market share prices with home market share prices. In particular, we identify the closing time for each home market from the StockMarketClock website, and compute the corresponding New York time.[12] The midpoint of the prevailing bid and ask quotes for the US cross-listed share is extracted from the Thomson Reuters Tick History (TRTH) database, and compared with the closing price of the home-market share. Consider an example of UK stocks cross-listed in the US. The home-market shares are traded on the London Stock Exchange between 9:00 am and 4:30 pm Greenwich Mean Time (GMT). Their cross-listed counterparts on the New York Stock Exchange are traded 5 hours behind GMT. To compare the synchronous prices of the UK-traded shares with the US-traded cross-listed shares, we obtain the midpoint of bid/ask quotes in the US market in effect at 11:30 am Eastern Standard Time (4:30 GMT). In cases where home and US markets are closed at the same time (e.g., Canada), we use the daily closing prices from both markets.

For those foreign markets which have completely non-overlapping trading hours with the US market (e.g., Australia and Japan), we use the midpoint of prevailing bid and ask quotes for US cross-listed shares within the first 15 minutes after the market opens to match with the home-market closing prices. We discard stocks with missing or invalid prices over the sample period.[13] The price deviation for a home-US pair is computed by taking the natural logarithm of its US price (P_t^{US}) expressed in US dollars divided by its home-market share price (P_t^H) also denominated in US dollars, and

[11] Canadian Stocks trade synchronously with their cross-listed counterparts in the US, but they are the exception.
[12] See https://www.stockmarketclock.com/exchanges.
[13] We detected unusually high bid/ask quotes for some stocks in the US markets and realized that this was due to recording error in TRTH. As a result, we set it as a missing value when such erroneous quotes were encountered.

adjusted for the ADR bundling ratio if relevant at any point in time.[14] The price deviation is computed as $\ln(P_{US}/P_H)$.

To examine the effect of investor attention comovement on price deviations in the cross-section of cross-listed stocks, we perform a panel regression across firm-quarters:

$$|\ln(P_{US}/P_H)|_{i,q} = a + \beta_1 AttentionComove_{i,q} + \sum_{k=2}^{K} \beta_k Control_{i,k,q} + \varepsilon_{i,q} \qquad (2.6)$$

where $|\ln(P_{US}/P_H)|_{i,q}$ is the absolute value of the average price deviations for firm i in quarter q, calculated by averaging daily price deviations across the days for which data are available in quarter q. $AttentionComove_{i,q}$ is the attention comovement obtained from Equation (2.4) in quarter q. $Control$ are a set of control variables, including idiosyncratic risk, dividend yield, interest rate, home-market volatility, US market volatility, FX volatility, market capitalization, home illiquidity, US illiquidity, institutional ownership, number of analysts and dispersion of analysts (See Appendix A2.2 for details of these specific variables). These variables are used to proxy for holding costs, transaction costs, and information-based barriers, which are found to be significantly related to price deviations in cross-listed stocks.

The primary interest from Equation (2.6) is β_1. Hypothesis 3 predicts that attention comovement is negatively associated with price deviations for cross-listed stock pairs. In addition, we expect price deviations to be significantly associated with holding costs, transaction costs, and information asymmetry.

[14] We convert the home-market share price to US dollars when extracting the data series from Datastream. Datastream applies the closing WMR (World Market Reuters) rate, which is normally updated by 4:00 PM GMT London Time, for the exchange in each particular date.

2.4.3 Summary statistics

Table 2.2 reports descriptive statistics for the key variables used in the analysis. Panel A presents the summary statistics for firm characteristics, and Panel B presents the summary statistics for market characteristics. On average, attention comovement in home-US cross-listed pairs is -1.4520, with an adjusted R^2 of 23.17% (from Equation (2.3)). This suggests that almost a quarter of the variation in investor attention can be explained by the within-pair counterparts. The mean market capitalization of firms in our sample is $16.06 billion, ranging from $1.2 million for the smallest firm to $152.13 billion for the largest firm. We report an average price difference of 65 basis points, which indicates that, US cross-listed shares trade at a modest premium relative to their home-market share counterparts. Our result is comparable to Gagnon and Karolyi (2010), who document a mean price difference of 32 basis points over the 1990 – 2004 period. The distribution of price differences is right skewed, with a minimum of 3.73% discount and a maximum of 14.25% premium. The medium price difference is 4 basis points. The average idiosyncratic risk of return difference is 1.29% per quarter, which is lower than 1.95% documented in Gagnon and Karolyi (2010). Across all firm-quarters for which data is available, the average dividend yield is 0.50% per quarter and it ranges from zero to a maximum of 5.71% per quarter. It is noteworthy that the median dividend yield is zero, which suggests the majority of the cross-listed firms in our sample did not pay dividends. Using the same dividend yield measure, Gagnon and Karolyi (2010) document an average dividend yield of 0.43% for their cross-listed sample.

The mean Amihud (2002) illiquidity ratio for the home-market and cross-listed shares are 0.4153 and 0.4008, respectively. According to the IBES file, 13 analysts cover a typical firm from its home market, on average, and six analysts cover its cross-listed counterpart in the US market. This result is similar to Gagnon and Karolyi (2010), who document the mean analyst coverage of 14 in the home market, and five in the US market. The analyst coverage also exhibits a significant variation in the cross-section of firms, with a standard deviation of 10 in the home markets and 6 in the US market. The mean value of the US institutional ownership is 19.26%, ranging from a minimum value barely distinguishable from zero to a maximum value of 99.76%. The mean earning announcement speed is -0.1533, indicating firms, on average, report their annual earnings approximately 56 days after their fiscal year end. The result is in accord with Securities & Exchange Commission (SEC)'s requirement that public companies need to file annual earnings reports no later than 60 days after their fiscal year end. [15]

[15] The 60-day filing requirements is set force by the SEC on 27 August, 2002. Before that, companies were allowed the 90-day requirements.

At the market level, Panel B shows that the average bank lending rate in the home market is 7.73% per annum. The mean value of the ratio of stock market capitalization to GDP (financial development) is 1.29, with a minimum value of 0.06 for the least developed market, and a maximum value of 12.54 for the most developed market. The average stock market turnover in the home country is 0.66 and it ranges from 0.02 for the most illiquid market to 4.08 for the most liquid market. In terms of legal protection, the average disclosure index is 0.63 out of 1 and the average anti-director rights index is 3.09 out of 5.

Panel C displays the cross-sectional correlations among the key firm characteristics. Since market value, price and book-to-market are logged in the cross-sectional regressions to follow, the correlations are also based on natural logs of these variables. Overall, the correlations among our key variables are low. Among all the variables, attention comovement is most correlated with US number of analysts (0.33), return volatility (0.23), and institutional ownership (0.22). We find that market value is negatively correlated with return volatility (-0.58), ROA volatility (-0.51), and idiosyncratic risk (-0.57), but positively correlated with home number of analysts (0.64). Panel C also demonstrates that idiosyncratic risk is highly correlated with return volatility (0.59), and stock illiquidity for both home-market shares (0.53) and US cross-listed shares (0.51).

Table 2.2 Summary statistics

This table presents descriptive statistics for key variables used in the study. Panel A reports summary statistics for firm characteristics. Attention R^2 is the adjusted R^2 from the time-series regression of volume shocks to home-market shares on volume shocks to US cross-listed shares as specified in Equation (2.3). AttentionComove represents attention comovement in the cross-listed pairs, which is computed by taking the logarithm transformation of Attention R^2 ($\ln(R^2/(1-R^2))$). Price difference is the natural logarithm of the US share price expressed in US dollar divided by the home-market share price also expressed in US dollar and adjusted for the ADR bundling ratio. Market value is the market capitalization measured in billions of dollars. Price is the home-market share price expressed in US dollars. Book-to-market is the book value of the common equity divided by the market value of the common equity. Return volatility is standard deviation of the stock returns for the home-market shares, and ROA volatility is standard deviation of return on assets. Idiosyncratic risk is the standard deviation of the residuals from Equation (A2.1). Dividend yield is the total dollar value of the dividends paid over the quarter divided by the price of the share at the end of the quarter for the US cross-listed shares. Home and US illiquidity are measured by Amihud (2002) illiquidity ratio. Number of analysts is the number of estimates underpinning the one-fiscal-year-ahead (FY1) EPS published in IBES. Dispersion of analysts is computed by dividing the standard deviation of each quarter's earnings forecasts outstanding across all analysts by the absolute value of the mean estimate across all analysts in the home market. Institutional ownership is the percentage of a firm's shares outstanding held by institutional investors. Earnings announcement speed is the number of days between the end of the fiscal year and the earnings announcement date, divided by 365 and multiplied by negative one. Panel B reports summary statistics for market characteristics. Interest rate is the home-market annual bank lending rate extracted from World Bank website. Home-market volatility is the standard deviation of returns on the home country index. US market volatility is the standard deviation of Standard and Poor's 500 index returns. Currency volatility is the standard deviation of foreign exchange rate for home countries. Financial development is the ratio of stock market capitalization of listed domestic companies to GDP. Stock market turnover is the ratio of the value of total shares traded to market capitalization. Anti-director rights index and Disclosure index are country-level indexes, obtained from La Porta, Lopez-de-Silanes, Shleifer, and Vishny (1998) and La Porta, Lopez-de-Silanes, and Shleifer (2006), respectively. Panel C reports correlations among key variables. Correlations are estimated in the cross section each quarter and then averaged over time. All continuous variables are winsorized at the 1 and 99 percentiles before estimating summary statistics and correlations.

	Mean	Std. Dev.	Min	P25	P50	P75	Max
Panel A:Firm characteristics							
Attention R^2	0.2317	0.1942	-0.0455	0.0678	0.1948	0.3616	0.9482
AttentionComove	-1.4520	1.3953	-10.8020	-2.1951	-1.2595	-0.4822	2.9062
Price difference	0.0065	0.0237	-0.0373	-0.0009	0.0004	0.0035	0.1425
Absolute price difference	0.0206	0.0672	0.0001	0.0007	0.0020	0.0067	0.3739
Market value ($ billions)	16.0629	39.1585	0.0012	0.4113	2.8300	15.4196	152.1346
Price	17.8111	27.7676	0.0800	2.8890	8.7200	21.8200	200.8200
Book-to-market	0.6536	0.5471	0.0014	0.3096	0.5102	0.8130	3.7037
Return volatility	0.0287	0.0170	0.0083	0.0171	0.0240	0.0351	0.1022
ROA volatility	0.0632	0.0785	0.0007	0.0151	0.0357	0.0786	0.4161
Idiosyncratic risk	0.0129	0.0128	0.0018	0.0049	0.0087	0.0158	0.1055
Dividend yield	0.0050	0.0102	0.0000	0.0000	0.0000	0.0060	0.0571
Home illiquidity	0.4153	1.3477	0.0000	0.0003	0.0031	0.0701	13.0668
US illiquidity	0.4008	1.2298	0.0000	0.0010	0.0075	0.1061	12.5771
Home number of analysts	13.0000	10.0000	1.0000	5.0000	12.0000	19.0000	59.0000
US number of analysts	6.0000	6.0000	1.0000	2.0000	4.0000	8.0000	47.0000
Dispersion of analysts	0.3052	0.6370	0.0000	0.0588	0.1173	0.2560	4.6000
Institutional ownership	0.1926	0.2212	0.0000	0.0161	0.0903	0.3233	0.9976
Announcement speed	-0.1533	0.0837	-0.5699	-0.1836	-0.1342	-0.0986	-0.0411
Panel B: Market characteristics							
Interest rate	0.0773	0.1028	0.0050	0.0300	0.0496	0.0691	0.8636
Home-market volatility	0.0132	0.0071	0.0044	0.0088	0.0113	0.0158	0.0816
US-market volatility	0.0106	0.0057	0.0046	0.0067	0.0088	0.0124	0.0418
FX volatility	0.0117	0.0117	0.0000	0.0028	0.0098	0.0171	0.1554
Financial development	1.2861	1.5776	0.0627	0.6713	1.0695	1.2929	12.5447
Stock market turnover	0.6553	0.4133	0.0163	0.4046	0.6289	0.7706	4.0788
Disclosure index	0.6253	0.1970	0.1700	0.5000	0.6250	0.7500	1.0000
Anti-director rights index	3.0882	1.3788	0.0000	2.0000	3.0000	4.0000	5.0000

Table 2.2 continued

Panel C: Correlation matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. AttentionComove	1														
2. Market value	-0.0345	1													
3. Price	0.0546	0.4106	1												
4. Book-to-market	-0.0416	-0.1477	-0.1916	1											
5. Return volatility	0.2278	-0.5781	-0.3137	0.0446	1										
6. ROA volatility	0.1377	-0.5097	-0.2034	-0.1277	0.4793	1									
7. Idiosyncratic risk	-0.0622	-0.5740	-0.4264	0.0727	0.5879	0.3681	1								
8. Dividend yield	-0.0801	0.2092	0.0333	0.0461	-0.2042	-0.1914	-0.1414	1							
9. Home illiquidity	-0.0450	-0.5147	-0.2667	0.0355	0.4860	0.3256	0.5336	-0.1145	1						
10. US illiquidity	-0.1143	-0.4733	-0.2432	0.0955	0.3730	0.2721	0.5119	-0.0881	0.4925	1					
11. Home number of analysts	-0.0224	0.6400	0.2653	-0.0969	-0.2761	-0.2718	-0.3299	0.0696	-0.2663	-0.2406	1				
12. US number of analysts	0.3298	0.1011	0.2216	-0.1426	-0.0038	0.1037	-0.2314	-0.0879	-0.0921	-0.1635	0.1461	1			
13. Dispersion of analysts	0.0502	-0.2338	-0.1585	0.1227	0.2379	0.1762	0.1658	-0.0956	0.1036	0.0985	-0.1367	-0.0503	1		
14. Institutional ownership	0.2182	-0.0346	0.3431	-0.0615	-0.0122	0.0403	-0.2281	-0.1198	-0.0598	-0.1502	-0.1266	0.5647	0.0261	1	
15. Announcement speed	0.0297	0.2179	0.2720	-0.0565	-0.1922	-0.1043	-0.1535	-0.0010	-0.0791	-0.1307	0.1738	0.2769	-0.1138	0.1778	1

2.5 Empirical results

2.5.1 Existence of attention comovement

We begin the empirical analysis with a preliminary investigation of the existence of investor attention comovement in home-US stock pairs. Specifically, we run Equations (2.2a) and (2.2b) within each firm-quarter over the sample period from October 1996 to December 2016. Table 2.3 presents the summary for the coefficients of interest - β_1^{US} and β_1^{H}. We report the mean, median, the number of firm-quarters evaluated (Firm-qtrs), and the number (N) and the proportion (Ratio) of firm-quarters for which the coefficients are positive and statistically significant at the 5% level. We also break these statistics down by region and year, and report them in Panels A and B, respectively.

The first 3 columns report the regression estimates from Equation (2.2a), and the next 3 columns report the results from Equation (2.2b). In Panel A, the sample wide mean coefficient is around 0.39 for both β_1^{US} and β_1^{HM}. Of the 26,820 firm-quarters, there are 17,970 significant positive (at the 5% level) $\beta_1 s$, accounting for 67% of the sample.[16] β_1^{US} and β_1^{H} capture the attention comovement between the home-US stock pairs. Therefore, the result reveals strong attention comovement in cross-listed stock pairs in our sample, providing supportive evidence for Hypothesis 1.

We summarize regression estimates for the control variables in Equations (2.2a) and (2.2b) in Appendix A2.1. Our results show positive attention autocorrelations in both home-market and cross-listed shares. Attention autocorrelation coefficient (β_2) is equal to 0.22 for home-market shares and 0.23 for cross-listed shares, suggesting a modest persistence in volume shock from day to day. Consistent with Drake et al. (2017), we document strong comovement in attention between individual stocks and the market where the stock is listed. The average coefficient for β_4 is 0.64 and 0.45 in Equations (2.2a) and (2.2b), respectively. Furthermore, β_5 equals 0.06 in Equation (2.2a) and 0.04 in Equation (2.2b). The number of firm-quarters for which β_5 is significantly positive at the 5% level is 1,827 in Equation (2.2a) and 1,393 in Equation (2.2b). Since β_5 captures the extent to which investor attention to a cross-listed stock is affected by contemporaneous aggregate attention shocks in its counterpart market, our finding suggests that market-wide attention shocks have asymmetric effects on the cross-listed pairs. There is a tendency that aggregate attention shocks in the US market have a stronger impact on the home-market shares, while aggregate attention shocks in the

[16] Since the number of firm-quarters for which the coefficient of β_1 is significantly positive is always same for regressions (2a) and (2b), we don't report them separately.

home market have less impact on the US cross-listed shares.[17] Finally, we also control for the lagged aggregate attention. Results on β_6 and β_7 suggest that the impact of lagged aggregated attention is dwarfed by the contemporaneous aggregated attention.

In Table 2.3, we also observe interesting patterns in attention comovement across regions and over time. In Panel A, Canadian firms cross listed in the US exhibit the strongest attention comovement. The average coefficients on β_1^{US} and β_1^{H} are 0.55 and 0.43, respectively, with 9,054 out of 11,053 firm-quarters (81.91%) significantly positive at the 5% level. We notice a monotonic decrease in the magnitude of the attention comovement from Canada, to Latin America, Europe, Africa and Middle East, and finally to Asia-Pacific. This is reflected in decreased regression coefficients, and the proportion of firm-quarters with significant positive coefficients. Asia-Pacific cross-listed stocks display the lowest level of attention comovement. The mean estimates of β_1^{US} and β_1^{HM} are 0.19 and 0.28, with only 39.31% of firm-quarters produce significant positive coefficients. The observed pattern across regions is consistent with the view that geographic and cultural proximity affects the information flows between the markets (Davis and Henderson, 2008; Sarkissian and Schill, 2004; Grinblatt and Keloharju, 2001).

Table 2.3 Panel B demonstrates an upward trend in attention comovement over the past two decades. There is a noticeable increase in coefficient estimates. The fraction of firm-quarters with significant positive coefficients increases dramatically from 38.24% in 1996 to 76.95% in 2016. This finding is likely to be attributable to the falling trend in market segmentation (Brooks and Del Negro, 2004; De Jong and De Roon, 2005; Kizys and Pierdzioch, 2009).

[17] This is not the case for stocks domiciled in Africa/Middle East and Asia-Pacific. For these stocks, the number of significant positive $\beta_{5,H}$ is larger than that on $\beta_{5,US}$, suggesting the impact of home-market attention shocks on US cross-listed shares is stronger than the impact of US-market attention shocks on home-market shares. This finding is mostly likely due to the fact that US market trades with large time lag from these two regions.

Table 2.3 Attention comovement in cross-listed stock pairs

This table shows the relation of investor attention between the home-market and cross-listed shares across 26,820 firm-quarters for 662 US cross-listed stocks included in our sample from 1996 to 2016. For each firm-quarter, we estimate the following time-series regressions:

$$V_{i,t}^{H} = a_{0i} + \beta_{1i}^{US} V_{i,t}^{US} + \beta_{2i} V_{i,t-1}^{H} + \beta_{3i} V_{i,t-1}^{US} + \beta_{4i,H} V_{H,t} + \beta_{5i,US} V_{US,t} + \beta_{6i,H} V_{H,t-1} + \beta_{7i,US} V_{US,t-1} + \varepsilon_{i,t} \quad (2.2a)$$

$$V_{i,t}^{US} = a_{0i} + \beta_{1i}^{H} V_{i,t}^{H} + \beta_{2i} V_{i,t-1}^{US} + \beta_{3i} V_{i,t-1}^{H} + \beta_{4i,US} V_{US,t} + \beta_{5i,H} V_{H,t} + \beta_{6i,US} V_{US,t-1} + \beta_{7i,H} V_{H,t-1} + \varepsilon_{i,t} \quad (2.2b)$$

where $V_{i,t}^{H}$ ($V_{i,t}^{US}$) is the volume shock for stock *i* on day *t* in the home market (the US market) as specified in Equation (2.1). We also include $V_{i,t-1}^{H}$ and $V_{i,t-1}^{US}$ in the regressions to control for autocorrelations in attention shocks, and include the lagged and contemporaneous aggregate volume shocks (V_H and V_{US}) to control for market-wide attention shocks. β_1^{US} and β_1^{H} captures attention comovement between the cross-listed pairs. For both β_1^{US} and β_1^{H}, we report the group mean and median across all firm-quarters, and the number (N) and proportion (Ratio) of positive coefficients that are statistically significant at the 5% level. We also report the coefficient estimates and associated statistics by home-market region and year in Panels A and B, respectively.

	β_1^{US}			β_1^{H}					
	Mean	Median	R^2	Mean	Median	R^2	Firm-qtrs.	N	Ratio
Panel A: By region									
All	0.3942	0.3533	0.36	0.3893	0.3800	0.33	26,820	17,970	67.00%
Canada	0.5459	0.5462	0.38	0.4267	0.4132	0.39	11,053	9,054	81.91%
Latin American	0.4314	0.4295	0.34	0.4162	0.4105	0.33	3,936	2,933	74.52%
Europe	0.2562	0.2375	0.37	0.3927	0.3898	0.29	6,151	3,519	57.21%
Africa/Middle East	0.2907	0.2674	0.33	0.3197	0.2865	0.26	1,864	964	51.72%
Asia-Pacific	0.1893	0.1734	0.35	0.2818	0.2777	0.25	3,816	1,500	39.31%
Panel B: By year									
1996	0.2374	0.1583	0.31	0.2815	0.2625	0.21	170	65	38.24%
1997	0.2732	0.2036	0.25	0.3021	0.2853	0.20	782	366	46.80%
1998	0.2923	0.2238	0.26	0.3215	0.2960	0.21	931	460	49.41%
1999	0.3044	0.2382	0.30	0.3611	0.3347	0.23	974	527	54.11%
2000	0.2923	0.2478	0.30	0.3735	0.3669	0.25	988	562	56.88%
2001	0.2804	0.2544	0.33	0.3648	0.3624	0.25	1,067	596	55.86%
2002	0.3406	0.2572	0.33	0.3583	0.3575	0.26	1,191	675	56.68%
2003	0.3277	0.2732	0.33	0.3727	0.3676	0.28	1,290	774	60.00%
2004	0.3882	0.3283	0.34	0.3803	0.3707	0.30	1,372	884	64.43%
2005	0.3960	0.3538	0.32	0.3799	0.3677	0.29	1,434	958	66.81%
2006	0.4198	0.3763	0.35	0.3798	0.3595	0.33	1,471	999	67.91%
2007	0.4362	0.4105	0.37	0.3900	0.3778	0.37	1,446	1,037	71.72%
2008	0.4346	0.4126	0.42	0.3773	0.3625	0.40	1,454	1,014	69.74%
2009	0.4275	0.3827	0.39	0.3953	0.3899	0.38	1,461	1,024	70.09%
2010	0.4423	0.4110	0.40	0.4064	0.4043	0.38	1,495	1,090	72.91%
2011	0.4477	0.4123	0.41	0.3930	0.3757	0.40	1,534	1,112	72.49%
2012	0.4441	0.4224	0.37	0.4059	0.3922	0.34	1,520	1,111	73.09%
2013	0.4359	0.4063	0.37	0.4134	0.4085	0.35	1,530	1,134	74.12%
2014	0.4446	0.4260	0.41	0.4290	0.4251	0.37	1,558	1,183	75.93%
2015	0.4289	0.4042	0.41	0.4280	0.4220	0.37	1,573	1,184	75.27%
2016	0.4246	0.4049	0.43	0.4580	0.4599	0.39	1,579	1,215	76.95%

Table 2.4 The determinants of attention comovement

This table presents the results from different specifications of Equation (2.5). The dependent variable is the attention comovement estimated from Equation (2.4). Earnings announcement speed is measured as the number of days between the end of the fiscal year and the earnings announcement date, divided by 365 and multiplied by negative one. Home and US number of analysts is based on the number of estimates underpinning the one-fiscal-year-ahead (FY1) EPS published in IBES international and US summary files. Home and US illiquidity is measured by Amihud (2002) illiquidity ratio, which is the average ratio of the daily absolute return to the dollar trading volume(in millions of dollars) over the number of days for which data are available during each quarter. Return volatility is the standard deviation of daily returns on the home-market shares over the quarter. ROA volatility is the standard deviation of the return on assets over the past 5 years. Home (US) market aggregate attention is the equal-weighted volume shocks to all the stocks in the home (US) market, excluding the firm in the measure of attention comovement. Market value is the natural logarithm of a firm's market capitalization expressed in US dollars. Price is the natural logarithm of home-market share price expressed in US dollars. Book-to-market is the book value of the common equity divided by the market value of the common equity and natural logged. Market integration is the correlation coefficient between returns on home-country index and returns on S&P's 500 index over the previous 60 months. Financial development is the ratio of stock market capitalization of listed domestic companies to GDP. Stock market turnover is the ratio of the value of total shares traded to market capitalization. Financial development and stock market turnover are obtained for each home market from the World Bank Website at the annual frequency. Anti-director rights index and Disclosure index are country-level indexes, obtained from La Porta et al. (1998) and La Porta et al. (2006), respectively. Column P.Sign indicates the predicted sign of each variable. The table reports the regression results across firm-quarters with country and quarter fixed effects for Models (1) to (5), and with regional fixed effects for Model (6). T-statistics based on firm-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	P.Sign	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Earnings announcement speed	+	0.8060***		0.6628**	0.6537**		0.9560***
		(3.11)		(2.36)	(2.31)		(3.56)
Home number of analysts	+	0.0047		0.0074**	0.0092***		0.0124***
		(1.61)		(2.53)	(3.21)		(3.83)
US number of analysts	+	0.0270***		0.0230***	0.0237***		0.0120*
		(4.61)		(3.79)	(4.04)		(1.86)
Home illiquidity	-	-0.1762		-0.2045	-0.2091		-0.3407**
		(-1.46)		(-1.46)	(-1.52)		(-2.57)
US illiquidity	-	-0.1324*		-0.1300*	-0.1227*		-0.1167
		(-1.82)		(-1.91)	(-1.73)		(-1.60)
Institutional ownership	+	0.4288*		0.3602**	0.4083**		0.2661
		(1.92)		(2.00)	(2.31)		(1.08)
Return volatility	+		22.7654***	26.7156***	25.6814***		23.2153***
			(20.29)	(11.18)	(10.83)		(9.64)
ROA volatility	+		0.8347***	2.7351***	2.6749***		2.8957***
			(3.17)	(4.53)	(4.42)		(4.33)
Home market aggregate attention	+				0.2265***	0.0733	0.0459
					(3.06)	(1.25)	(0.59)

Table 2.4 continued

US market aggregate attention	+				0.6505***	0.6131***	0.3377*
					(2.75)	(5.07)	(1.77)
Market integration	+				-0.6487**	-0.2574	0.2679
					(-2.10)	(-1.64)	(1.58)
Market value	+	0.0040	0.1352***	0.0764***	0.0731***	0.0404***	0.0946***
		(0.18)	(10.45)	(3.20)	(3.05)	(3.65)	(3.82)
Price	+	0.0178	0.0363**	0.0945***	0.0888***	-0.0183	0.0922***
		(0.74)	(2.08)	(3.94)	(3.80)	(-1.24)	(4.50)
Book-to-market	+/-	0.0703**	-0.0202	0.0954***	0.1036***	-0.0397**	0.0936***
		(2.10)	(-0.96)	(2.78)	(3.05)	(-1.97)	(2.93)
Lag AttentionComove	+	0.2081***	0.2998***	0.1690***	0.1702***	0.3484***	0.2103***
		(12.58)	(22.49)	(10.87)	(10.97)	(26.46)	(11.61)
Financial development	+						0.0926***
							(8.06)
Stock market turnover	+						0.0574
							(1.05)
Anti-director rights index	+						-0.0418*
							(-1.92)
Disclosure index	+						0.3568
							(1.48)
Constant		-1.8772***	-3.0422***	-3.2140***	-2.9564***	-1.7493***	-2.9545***
		(-7.27)	(-20.42)	(-10.59)	(-10.06)	(-12.00)	(-7.33)
Country fixed-effects		Yes	Yes	Yes	Yes	Yes	No
Regional fixed-effects		No	No	No	No	No	Yes
Quarter fixed-effects		Yes	Yes	Yes	Yes	Yes	No
Observations		5,942	17,095	5,661	5,616	18,710	4,644
R-squared		0.2670	0.3122	0.3092	0.3116	0.2668	0.2693

2.5.2 The determinants of attention comovement

Our results thus far demonstrate strong attention comovement between the home-market and US cross-listed shares. We now turn to examining the determinants of attention comovement. The objective of the analysis is to better understand how attention comovement is affected by information- and social-related factors as documented in previous studies. For this purpose, we turn our attention to the R^2 measure from Equation (2.4), denoted as *AttenionComove*, and use it as the measure for attention comovement in the subsequent analyses.[18] This provides a consistent measure which can be applied to different empirical settings. As discussed in Section 2.4.2.2, we include a set of firm and country characteristics that potentially affect attention comovement in cross-listed stock pairs. We run different specifications of Equation (2.5) across all firm-quarters using a panel regression approach, and report the results in Table 2.4. Models (1) to (5) present the regression estimates with country and quarter fixed-effects. Model (6) presents the estimates with regional fixed-effects. T-statistics are based on firm-clustered standard errors.[19]

In Model (1), we first regress *AttentionComove* on firm-level information environment variables controlling for size, price, book-to-market and lagged attention comovement. The signs of the coefficients are consistent with our expectations. The coefficient on earnings announcement speed is 0.8060 and is statistically significant (t = 3.11). The coefficient is positive for the number of analysts in both home and US markets (0.0047 and 0.0270, respectively), and is statistically significant for the number of analysts in the US (t = 4.61). We also observe negative relations between attention comovement and illiquidity in both home and US markets. In addition, attention comovement is positively associated with institutional ownership. For the control variables, both book-to-market and lagged attention comovement are significant. The significance on the lagged attention comovement variable is not surprising, as the linkage between cross-listed and home-market stock pairs should be correlated through time. Overall, results from Model (1) are consistent with the prediction that attention comovement is stronger for cross-listed firms that operate under a better information environment.

[18] The results from Table 2.3 support the use of R^2 measure. First, the contemporaneous relations are not affected when market and lagged attention variables are included. Second, Equations (2.2a) and (2.2b) give similar results on regression coefficient and R^2.

[19] Petersen (2009) advocates standard errors clustered by firm as unbiased with correctly sized confidence intervals for permanent or temporary firm effects when the residuals are correlated across firms or time.

In Model (2), we examine whether attention comovement is associated with information shocks as proxied by stock return volatility and ROA volatility. The result shows that coefficients of both variables are positive and statistically significant. Thus, it supports our conjecture that frequent information shocks draw intensive investor attention and therefore lead to stronger attention comovement. Model (3) combines Models (1) and (2) together, and the findings on the key variables remain robust. In particular, attention comovement is positively related to earnings announcement speed, number of analysts, institutional ownership, return volatility, ROA volatility, and is negatively related to stock illiquidity.

In Model (4), we augment Model (3) with social-related factors, including home market aggregate attention, US market aggregate attention, and market integration between home country and the US. Consistent with the behavioural literature which argues that investor attention is at least partially socially driven, our result reveals a strong positive relation between attention comovement and aggregate investor attention in the home and US markets. The magnitude of the coefficient on aggregate attention from the US market is much larger than that from the home market (0.6505 vs 0.2265). This indicates that the US market is more dominant in driving cross-market attention spillover. In particular, the aggregate attention shocks from the US markets could spill over to the home markets, resulting in stronger firm level attention comovement.

In Model (5), we focus on the relation between attention comovement and social-related factors. The result shows a significant positive coefficient on US market aggregate attention. However, we do not observe a significant relation between attention comovement and home market aggregate attention. Thus, the result further supports that aggregate attention in the US market is the key driver of attention comovement among cross-listed pairs. In Model (5), market integration exhibits no explanatory power on attention comovement.

Model (6) contains all the information- and social-related factors. It augments Model (4) by including variables associated with information environment at country level. The results indicate that the coefficient on financial development is positive and statistically significant (coefficient = 0.0926, t = 8.06). However, we document a marginal negative relation between attention comovement and Anti-director right index. Consistent with Model (4), our findings show that stocks with strong firm-specific information environment and frequent information shocks exhibit high attention comovement. In general, our empirical results tend to suggest that cross-sectional variation in attention comovement is better explained by firm-level information environment than by country-level information environment.

To summarize, there are three key findings from Table 2.4. First, home-US stock pairs with a high quality of information environment exhibit stronger attention comovement, and cross-listed firms that experience frequent information shocks are also associated with stronger attention comovement. Second, aggregate investor attention in the US market has a reliable positive effect on attention comovement. Finally, attention comovement variations are more affected by firm-specific information environment than by home country information environment. Overall, our results provide supportive evidence for both information-driven and socially-driven explanations for attention comovement.

2.5.3 Attention comovement and deviations from price parity

2.5.3.1 Attention comovement and contemporaneous price deviation

In this section, we examine the relation between attention comovement and deviations from price parity in cross-listed stock pairs. We investigate whether attention comovement incrementally explains variations in price deviations after controlling for numerous factors that are identified in prior studies as important attributes affecting ADR mispricing. Our test is conducted using a panel regression across firms and quarters. We run different specifications of Equation (2.6) with country and quarter fixed effects. Coefficients and t-statistics based on firm-clustered standard errors are presented in Table 2.5.

In the baseline model (Model (1)), the coefficient on *AttentionComove* is negative (-0.0029) and statistically significant (t = 3.66). The result suggests that cross-listed pairs with strong attention comovement are associated with smaller price deviations. In terms of economic magnitude of the relationship, a one-standard-deviation increase in attention comovement (standard deviation equals 1.40 per quarter) is associated with a 0.4% decrease in the absolute price disparity, which corresponds to about 17.13% of its standard deviation across all firm-quarters (2.37% from Panel A of Table 2.2).[20] The significant negative relation between attention comovement and price disparity remains robust after controlling for holding costs, transaction costs, and market volatility in Models (2) to (4). Coefficients on *AttentionComove* equal -0.0021, -0.0024 and -0.0026 in Models (2), (3) and (4), respectively, and are all significant at the 1% level.

[20] A 0.4% decrease in the absolute price disparity corresponds to 19.71% of the average absolute price disparity across all firm-quarters (the average absolute price disparity is 0.0206 from Panel A of Table 2.2).

The signs of control variables are generally consistent with the expectations and prior studies. For example, price deviations are positively related to idiosyncratic risk and lending interest rate (De Jong, Rosenthal, and Van Dijk, 2009; Gagnon and Karolyi, 2010; Grossmann et al., 2007). In addition, we document a significant positive relation between price disparity and home-market volatility. However, our result shows a negative relation between price disparity and US-market volatility.[21]

In Model (5), we augment Model (4) with variables associated with information barriers. The coefficient on *AttentionComove* remains statistically negative. By contrast, the result does not indicate that price disparity is related to institutional ownership, number of analysts, and dispersion of analysts, which are used to capture information barriers in prior studies.[22] This implies that attention comovement used in this study might be a better proxy for information flows compared to these conventional measures. Model (6) presents the result excluding *AttentionComove*. The coefficient on each variable is comparable to that in Model (5). Price deviations are significantly related to holding costs and market volatility, but not with information-based variables.

Overall, empirical results in Table 2.5 show a significant negative contemporaneous relation between attention comovement and price deviations in cross-listed pairs. The result is robust after controlling for a set of variables that are potential determinants of price deviations. Our finding is consistent with the argument that attention comovement captures the correlated information flows between the paired shares; strong attention comovement reduces information barriers and as a result, reduces the price deviation. It is worth noting that, attention comovement is the only variable that is significantly associated with price deviations among those proxies for information barriers. A possible explanation is that attention comovement is a bilateral measure capturing the correlated information flows from home and US markets, while institutional ownership, number of analysts, and dispersion of analysts are all unilateral measures of information barriers.

[21] We investigate the possible reasons for this counterintuitive finding by excluding home-market volatility from the regression. We document a positive but insignificant relation between price disparity and US-market volatility. Thus, the observed negative relation between price disparity and US market volatility is most likely driven by a high correlation between home-market and US-market volatility (0.76).
[22] We also run a separate regression on those three variables. The result shows insignificant coefficients on institutional ownership, number of analysts and analyst dispersion.

Table 2.5 Attention comovement and price deviation

The table reports the results from different specifications of Equation (2.6). The dependent variable is the absolute price deviation of a home-US stock pair, measured by averaging $Ln(P_{US}/P_H)_t$ of each day across a quarter and taking its absolute value. AttentionComove is the attention comovement measure specified in Equation (2.4). Idiosyncratic risk is the standard deviation of the residuals from Equation (A2.1). Dividend yield is computed by dividing the total dollar value of the dividend paid over the quarter by the share price at the end of the quarter for the US cross-listed shares. Interest rate is the home-market annual lending interest rate. Market value is the natural logarithm of a firm's market capitalization expressed in US dollars. Home-market volatility is the standard deviation of returns on the home market index. US market volatility is the standard deviation of S&P'S 500 index returns. Currency volatility is the standard deviation of foreign exchange rate for the home country. Home and US illiquidity is measured by Amihud (2002) illiquidity ratios. Institutional ownership is the percentage of a firm's shares outstanding held by institutional investors. Number of analysts is the number of estimates underpinning the one-fiscal-year-ahead (FY1) EPS published in IBES international file. Dispersion of analysts is computed by dividing the standard deviation of each quarter's earnings forecasts outstanding across all analysts by the absolute value of the mean estimate across all analysts. We report the panel regression results across firms and quarters with country and quarter fixed-effects. T-statistics are based on firm-clustered standard errors. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
AttentionComove	-0.0029***	-0.0021***	-0.0024***	-0.0026***	-0.0027***	
	(-3.66)	(-2.60)	(-3.00)	(-3.26)	(-2.75)	
Idiosyncratic risk		0.5481***	0.5813**	0.5410**	0.6927**	0.6731**
		(2.83)	(2.58)	(2.40)	(2.12)	(2.08)
Dividend yield		-0.0824	-0.0625	-0.0577	0.0846	0.0963
		(-0.71)	(-0.62)	(-0.58)	(0.63)	(0.74)
Interest rate		0.1923***	0.1881***	0.1564**	0.1509***	0.1608***
		(2.95)	(2.87)	(2.45)	(2.60)	(2.76)
Market value			-0.0011	-0.0012	0.0006	0.0008
			(-0.79)	(-0.83)	(0.34)	(0.45)
Home illiquidity			0.0004	0.0005	-0.0017	-0.0015
			(0.17)	(0.20)	(-0.92)	(-0.84)
US illiquidity			-0.0035	-0.0034	-0.0033**	-0.0026**
			(-1.52)	(-1.49)	(-2.40)	(-2.02)
Home-market volatility				1.6028***	1.4953***	1.3909**
				(3.67)	(2.75)	(2.57)
US-market volatility				-1.3721***	-1.3963***	-1.3058***
				(-3.38)	(-2.77)	(-2.60)
Currency volatility				-0.0563	-0.0448	-0.0570
				(-0.59)	(-0.34)	(-0.44)
US institutional ownership					-0.0078	-0.0096
					(-1.07)	(-1.35)
Home number of analysts					-0.0005	-0.0005
					(-1.15)	(-1.24)
Home dispersion of analysts					0.0000	-0.0001
					(0.01)	(-0.05)
Constant	0.1429***	0.0461**	0.0536**	0.0466*	0.0454*	0.0478*
	(5.37)	(2.21)	(2.17)	(1.86)	(1.70)	(1.83)
Country fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Quarter fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,192	19,296	19,269	19,269	9,738	9,913
R-squared	0.1327	0.1789	0.1811	0.1871	0.2286	0.2221

2.5.3.2 Unilateral investor attention and price deviation

Empirical results from Table 2.5 show that correlated investor attention enhances price parity in home-US stock pairs. In this section, we examine how unilateral investor attention affects deviations from price parity. In particular, we investigate the impact of attention from either home or US market on price deviations among cross-listed pairs. Prior studies suggest that cross-border information imbalance affects the relative pricing of cross-listed pairs (e.g., Chen and Choi, 2012). We conjecture that high investor attention arises from unilateral markets may lead to a larger cross-border information imbalance, and as a result, a larger price deviation. To test this conjecture, we repeat the same analysis as in the previous section but include firm-specific attention from either home or US market instead of attention comovement. The results are shown in Table 2.6.

Model (1) presents the relation between price deviations and US-market investor attention. The coefficient on US attention is positive (0.0042) and statistically significant at the 5% level (t = 2.29). This suggests that higher US investor attention leads to larger deviations from price parity for the cross-listed pairs. Model (2) presents the relation between price deviations and home-market investor attention. The results show that home-market investor attention has a positive effect on price disparity, although the impact is not statistically significant. In Model (3), we include both US- and home-market attention in the regression, and the result shows that increased investor attention in US market leads to larger price deviations, while increased investor attention in home market has no significant effect on price deviations.

Overall, results in Table 2.6 show that unilateral investor attention results in larger price deviations, and attention from US market plays a dominant role affecting price deviations in cross-listed pairs. Linking to the findings in Table 2.5, our results provide empirical evidence that correlated information flows leads to comovement in asset prices, while information imbalance between distinct trading locations can yield transitory disparities in the prices of an identical security (Hong and Stein, 1999; Chowdhry and Nanda, 1991).

Table 2.6 Unilateral investor attention and price deviation

The table reports the price deviations associated with US- and home-market attention. The dependent variable is the absolute price deviation of a home-US stock pair, measured by averaging $Ln(P_{US}/P_H)_t$ of each day across a quarter and taking its absolute value. US (Home) attention is the average of the daily volumes shocks to the US cross-listed (home-market) shares in each quarter. Idiosyncratic risk is the standard deviation of the residuals from Equation (A2.1). Dividend yield is computed by dividing the total dollar value of dividend paid over the quarter by the share price at the end of the quarter for the US cross-listed shares. Interest rate is the home-market annual lending interest rate. Market value is the natural logarithm of a firm's market capitalization expressed in US dollars. Home-market volatility is the standard deviation of returns on home market returns. US market volatility is the standard deviation of S&P'S 500 index returns. Currency volatility is the standard deviation of foreign exchange rate for the home country. Home and US illiquidity is measured by Amihud (2002) illiquidity ratios. Institutional ownership is the percentage of a firm's shares outstanding held by institutional investors. Number of analysts is the number of estimates underpinning the one-fiscal-year-ahead (FY1) EPS published in IBES international file. Dispersion of analysts is computed by dividing the standard deviation of each quarter's earnings forecasts outstanding across all analysts by the absolute value of the mean estimate across all analysts in the home market. We report the panel regression results across firms and quarters with country and quarter fixed-effects. T-statistics are based on firm-clustered standard errors. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	Model (1)	Model (2)	Model (3)
US attention	0.0042**		0.0052**
	(2.29)		(2.16)
Home attention		0.0012	-0.0022
		(0.52)	(-0.71)
Idiosyncratic risk	0.6622**	0.6718**	0.6621**
	(2.06)	(2.08)	(2.07)
Dividend yield	0.0956	0.0962	0.0956
	(0.73)	(0.74)	(0.73)
Interest rate	0.1631***	0.1615***	0.1625***
	(2.80)	(2.75)	(2.78)
Market value	0.0008	0.0008	0.0008
	(0.46)	(0.45)	(0.46)
Home illiquidity	-0.0014	-0.0014	-0.0016
	(-0.82)	(-0.78)	(-0.86)
US illiquidity	-0.0024*	-0.0026**	-0.0024*
	(-1.92)	(-2.03)	(-1.90)
Home-market volatility	1.3601**	1.3854**	1.3632**
	(2.51)	(2.56)	(2.51)
US-market volatility	-1.3104***	-1.3089***	-1.3059***
	(-2.61)	(-2.61)	(-2.60)
Currency volatility	-0.0593	-0.0571	-0.0597
	(-0.46)	(-0.44)	(-0.47)
US institutional ownership	-0.0097	-0.0097	-0.0097
	(-1.36)	(-1.35)	(-1.35)
Home number of analysts	-0.0005	-0.0005	-0.0005
	(-1.24)	(-1.24)	(-1.24)
Home dispersion of analysts	-0.0001	-0.0001	-0.0001
	(-0.04)	(-0.04)	(-0.04)
Constant	0.0467*	0.0476*	0.0468*
	(1.78)	(1.82)	(1.78)
Country fixed-effects	Yes	Yes	Yes
Quarter fixed-effects	Yes	Yes	Yes
Observations	9,913	9,913	9,913
R-squared	0.2226	0.2221	0.2227

2.5.3.3 Lead-lag effect in attention comovement

In this section, we examine the lead-lag relation between attention comovement and price deviation. In Section 2.5.3.1, we show that attention comovement reduces price deviations among home-US stock pairs. However, it may be argued that price deviations can also lead to decreased attention comovement. Prior research on the price discovery process (the incorporation of new information) for cross-listed stocks has shown that when there is a price deviation from the equilibrium, both home and cross-listed markets mutually adjust to restore the equality. However, the information share can vary significantly between the two markets.[23] Some studies find that the price discovery mostly take place in the home market (e.g., Bacidore and Sofianos, 2002; Gramming et al., 2005), while some document a larger proportion of price discovery in the cross-listed market (Eun and Sabherwal, 2003). In either case, it implies variation in investor attention between the two markets. From this perspective, price deviations can result in lower attention comovement in cross-listed pairs. As a result, attention comovement and price deviations are likely jointly determined.

Although we cannot perfectly isolate the causality between attention comovement and price deviations, we can test whether attention comovement is incrementally predictive of price deviations. To do so, we first regress price deviations on the lagged attention comovement with the same controls and fixed effects as employed in Models (5) of Table 2.5. Model (1) in Table 2.7 shows that the lagged attention comovement is negative and statistically significant (coefficient = -0.0025, t = 2.71), which suggests that attention comovement can help predict the price deviations over the following period. In Model (2), we regress price deviations on both lagged attention comovement and lagged price deviations. The negative relation between lagged attention comovement and price deviations remains robust. However, the magnitude of the coefficient is much smaller than in Model (1). Specifically, the coefficient on attention comovement in Model (1) is -0.0025, while it drops to -0.0004 in Model (2).

Overall, our investigation of the lead-lag relation between attention comovement and price deviation shows that lagged attention comovement is negatively associated with current-period price deviations after controlling for the lagged price deviations. The bottom line of this finding is that attention comovement is not merely the "by-product" of price deviations, otherwise, the effect of attention comovement would be overshadowed by the lagged price deviations. Thus, our results from the lead-

[23] In Hasbrouck (1995), a market's contribution to price discovery is its information share, defined as the proportion of the efficient price innovation variance that can be attributed to that market.

lag regression suggest that attention comovement plays an incremental role in predicting price deviations.

Table 2.7 Lagged attention comovement and price deviation

The table reports the lead-lag relation between attention comovement and price deviation. The dependent variable is the absolute price deviation in a cross-listed pair ($|Ln(P_{US}/P_H)|$), measured by averaging $Ln(P_{US}/P_H)_t$ of each day across a quarter and taking its absolute value. AttentionComove is the attention comovement measure specified in Equation (2.4). Lag AttentionComove and Lag $|Ln(P_{US}/P_H)|$ represent the attention comovement over the previous quarter and the price deviation over the previous quarter, respectively. All control variables are defined in Equation (2.6). We report the panel regression results across all firm-quarters with country and quarter fixed-effects. T-statistics are based on firm-clustered standard errors. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	Model (1)	Model (2)
Lag AttentionComove	-0.0025***	-0.0004**
	(-2.71)	(-2.40)
Lag $\lvert Ln(P_{US}/P_H)\rvert$		0.9147***
		(40.77)
Idiosyncratic risk	1.2017***	0.3388**
	(3.46)	(2.50)
Dividend yield	0.0823	0.0341
	(0.65)	(1.07)
Interest rate	0.1643***	-0.0406**
	(2.64)	(-2.57)
Market value	0.0013	0.0002
	(0.90)	(0.67)
Home illiquidity	-0.0035**	-0.0007
	(-2.34)	(-1.48)
US illiquidity	-0.0040***	-0.0008
	(-2.83)	(-1.44)
Home-market volatility	1.3744**	0.2467*
	(2.35)	(1.86)
US-market volatility	-1.6259***	-0.3444**
	(-2.99)	(-2.40)
Currency volatility	-0.0551	0.0153
	(-0.44)	(0.38)
US institutional ownership	-0.0087	-0.0012
	(-1.20)	(-1.14)
Home number of analysts	-0.0003	0.0000
	(-0.96)	(1.18)
Home dispersion of analysts	-0.0002	0.0001
	(-0.12)	(0.26)
Constant	0.0405*	0.0081*
	(1.67)	(1.73)
Country fixed-effects	Yes	Yes
Quarter fixed-effects	Yes	Yes
Observations	8,538	8,505
R-squared	0.2370	0.9010

2.6 Robustness tests

To test the robustness of the results, we first investigate how attention comovement varies with firm news announcements. We then use different detrending days to calculate volume shocks and attention comovement. Finally, we examine the relation between attention comovement and price deviations in subsamples and subperiods.

2.6.1. Attention comovement and news announcement

Time-series analysis in Section 2.5.1 reveals the existence of attention comovement in home-US stock pairs. Prior studies suggest that investor attention increases when new information is arriving (Sims, 2003; Peng, 2005; Drake at al., 2012, 2015). Using earnings announcement and dividend declaration as indicators for arrival of firm news, we investigate whether attention comovement in the cross-listed pairs increases around the news announcement.

We collect quarterly earnings announcement date and dividend declaration date from IBES and CRSP, and set the earnings announcement dummy (*Dearning)* and dividend declaration dummy (*Ddividend*) to 1 if date t falls into the 3-day window around earnings announcement and dividend declaration. A total of 342 firms in our sample have available data for the two news events over the sample period.[24] On average, each firm has 33 earnings announcements and 19 dividend declarations. For each of the 342 firms, we estimate the following time-series regressions based on daily observations over the sample period from October 1996 to December 2016:

$$V_{i,t}^{H} = a_i + \beta_{1i}^{US} V_{i,t}^{US} + \beta_{2i}^{US} V_{i,t}^{US} * Dearning_{i,t} + \beta_{3i}^{US} V_{i,t}^{US} * Ddividend_{i,t} + \beta_{4i} Dearning_{i,t} + \beta_{5i} Ddividend_{i,t} + \beta_{6i} V_{i,t-1}^{H} + \beta_{7i} V_{i,t-1}^{US} + \beta_{8i,H} V_{H,t} + \beta_{9i,US} V_{US,t} + \beta_{10i,H} V_{H,t-1} + \beta_{11i,US} V_{US,t-1} + \varepsilon_{i,t} \tag{2.7a}$$

$$V_{i,t}^{US} = a_i + \beta_{1i}^{H} V_{i,t}^{H} + \beta_{2i}^{H} V_{i,t}^{H} * Dearning_{i,t} + \beta_{3i}^{H} V_{i,t}^{H} * Ddividend_{i,t} + \beta_{4i} Dearning_{i,t} + \beta_{5i} Ddividend_{i,t} + \beta_{6i} V_{i,t-1}^{US} + \beta_{7i} V_{i,t-1}^{H} + \beta_{8i,US} V_{US,t} + \beta_{9i,H} V_{H,t} + \beta_{10i,US} V_{US,t-1} + \beta_{11i,H} V_{H,t-1} + \varepsilon_{i,t} \tag{2.7b}$$

where coefficients on the interaction terms, β_2 and β_3 capture the effect of news announcement on attention comovement in the cross-listed pairs. We predict attention comovement to be stronger during news announcement periods. Therefore, β_2 and β_3 are expected to be significantly positive across firms.

[24] We require companies to have both earnings announcement and dividend declaration over the sample period to ensure an equal number of observations on each coefficient. Our findings still hold without this requirement.

Table 2.8 presents the average coefficient estimates across firms and the associated statistics. For each coefficient, we report the mean, the number of firms for which the coefficient has positive values, and the number of firms for which the coefficient is statistically positive at the 5% level. We are most interested in β_1, β_2 and β_3. β_1 provides insights on whether attention comovement remains after controlling for news events. β_2 and β_3 capture the relative importance of the news item on attention comovement.

Panel A presents the regression results from Equation (2.7a). Consistent with the finding in Table 2.3, the mean coefficient of β_1^{US} remains positive (0.2976), with 98.25% of firms significant at the 5% level. The result shows that attention comovement in the cross-listed remains after controlling for news announcement events. The average coefficient of β_2^{US} for the attention-earnings announcement interaction variable is 0.0587. We see that β_2^{US} is reliably positive as it arises for the majority of firms (238 out of 342), among which 86 firms (25.15%) are significant at the 5% level. This result suggests that attention comovement in cross-listed pairs increases over the earnings announcement periods. Similarly, we observe a positive mean of 0.0627 for the attention-dividend announcement interaction (β_3^{US}), which provides the evidence that attention comovement is stronger around the firms' dividend declaration.

In Panel B, similar regression results are observed from Equation (2.7b). Coefficient of β_1^{H} is positive and statistically significant for 336 out of 342 firms. On average, the coefficient of β_2^{H} for attention-earnings announcement interaction is positive (0.0275), with 16.37% of the sample that are statistically significant. This provides further support for the prediction that attention comovement is stronger over the earnings announcement periods.

Overall, empirical findings in Table 2.8 confirm the existence of attention comovement after controlling for news announcement events. Also, our results show that attention comovement is stronger during news announcement periods, and this is most apparent around earnings announcements.

Table 2.8 News announcement and attention comovement in cross-listed pairs

This table shows the impact of company news announcement on the attention comovement between the home-market and the US cross-listed shares. For each firm, we estimate the following time-series regression over the entire sample period from 1996 to 2016:

$$V_{i,t}^{H} = a_i + \beta_{1i}^{US}V_{i,t}^{US} + \beta_{2i}^{US}V_{i,t}^{US} * Dearning_{i,t} + \beta_{3i}^{US}V_{i,t}^{US} * Ddividend_{i,t} + \beta_{4i}Dearning_{i,t} + \beta_{5i}Ddividend_{i,t} + \beta_{6i}V_{i,t-1}^{H} + \beta_{7i}V_{i,t-1}^{US} + \beta_{8i,H}V_{H,t} + \beta_{9i,US}V_{US,t} + \beta_{10i,H}V_{H,t-1} + \beta_{11i,US}V_{US,t-1} + \varepsilon_{i,t} \quad (2.7a)$$

$$V_{i,t}^{US} = a_i + \beta_{1i}^{H}V_{i,t}^{H} + \beta_{2i}^{H}V_{i,t}^{H} * Dearning_{i,t} + \beta_{3i}^{H}V_{i,t}^{H} * Ddividend_{i,t} + \beta_{4i}Dearning_{i,t} + \beta_{5i}Ddividend_{i,t} + \beta_{6i}V_{i,t-1}^{US} + \beta_{7i}V_{i,t-1}^{H} + \beta_{8i,US}V_{US,t} + \beta_{9i,H}V_{H,t} + \beta_{10i,US}V_{US,t-1} + \beta_{11i,H}V_{H,t-1} + \varepsilon_{i,t} \quad (2.7b)$$

where $V_{i,t}^{H}$ ($V_{i,t}^{US}$) is the volume shock for stock *i* on day *t* in the home market (the US market), and is calculated based on a detrended measure on the stock's log-turnover as specified in Equation (2.1). *Dearning* is a dummy variable that takes the value of one if date *t* falls into the three-day window [-1, 1] around the earnings announcement. *Ddividend* is a dummy variable that takes the value of one if date *t* falls into the three-day window [-1, 1] around the dividend declaration. We also include $V_{i,t-1}^{US}$ and $V_{i,t-1}^{H}$ in the regressions to control for the stock's volume shock autocorrelation, and include V_H and V_{US} to control for market-wide volume shocks. Coefficients on the interaction terms, β_2 and β_3 capture the effect of news announcement on attention comovement within the cross-listed pairs. For each coefficient estimate, we report the group mean across all firms, the associated t-statistics, the number of firms with positive coefficients, and the number of firms with positive coefficients that are statistically significant at the 5% level.

Panel A: Estimates from Equation (2.7a)												
	β_1^{US}	β_2^{US}	β_3^{US}	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	R^2
Mean	0.2976	0.0587	0.0627	0.0928	-0.0017	0.3986	-0.0518	0.6906	0.0643	-0.4326	-0.0281	0.3949
No of positive coeff.	342	238	198	286	184	342	103	339	220	10	133	
No of signif positive coeff.	336	86	29	134	11	340	59	331	125	0	34	
Firms	342	342	342	342	342	342	342	342	342	342	342	
Ratio	98.25%	25.15%	8.48%	39.18%	3.22%	99.42%	17.25%	96.78%	36.55%	0.00%	9.94%	
Panel B: Estimates from Equation (2.7b)												
	β_1^{H}	β_2^{H}	β_3^{H}	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	R^2
Mean	0.3467	0.0275	-0.0382	0.0587	0.0356	0.4238	-0.0833	0.5162	-0.0059	-0.2141	-0.0177	0.3734
No of positive coeff.	342	228	200	263	204	342	47	338	157	9	142	
No of signif positive coeff.	336	56	21	67	16	338	8	312	51	0	29	
Firms	342	342	342	342	342	342	342	342	342	342	342	
Ratio	98.25%	16.37%	6.14%	19.59%	4.68%	98.83%	2.34%	91.23%	14.91%	0.00%	8.48%	

2.6.2 Alternative definitions for volume shocks

In the main analysis, we measure volume shocks as logarithm of daily turnover detrended by subtracting its 200-day moving average. To ensure that our results are not influenced by the length of the detrending window, we run Equation (2.6) when attention is calculated based on 50-day and 100-day moving windows, respectively. The results are displayed in Table 2.9. Model (1) presents the results based on the 100-day estimation window, and Model (2) presents the results based on the 50-day window.

In Table 2.9, we document a significant negative relation between price disparity and attention comovement. After including all the control variables, coefficients on *AttentionComove* are -0.0017 and -0.0021, respectively, for the 100- and 50-day estimation windows, and both are significant at the 5% level. This result is comparable to the 200-day detrended measure reported in Table 2.5 (coefficient = -0.0027 and t = 2.75). Therefore, our results are not influenced by the detrending window employed when computing volume shocks.

Table 2.9 Attention comovement and price deviation - Alternative detrending windows

The table reports the results from Equation (2.6) when AttentionComove is estimated using different time windows. The dependent variable is the absolute price deviation in a cross-listed pair, measured by averaging $Ln(P_{US}/P_H)_t$ of each day across a quarter and taking its absolute value. AttentionComove is the log transformation of adjusted R^2 as specified in Equation (2.4). The adjusted R^2 is obtained by regressing volume shocks of the home-market share on the volume shocks of the US cross-listed share, where volume shocks is computed by subtracting a 100-day or 50-day moving average from daily turnover. AttentionComove in Model (1) is calculated based on a 100-day estimation window and AttentionComove in Model (2) is calculated based on a 50-day estimation window. All the control variables are defined in Equation (2.6). We report the panel regression results across all firm and quarters with country and quarter fixed-effects. T-statistics are based on firm-clustered standard errors. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	Model (1)	Model (2)
AttentionComove	-0.0017**	-0.0021**
	(-2.25)	(-2.33)
Idiosyncratic risk	0.6914**	0.6890**
	(2.14)	(2.14)
Dividend yield	0.0890	0.0763
	(0.68)	(0.58)
Interest rate	0.1467**	0.1601***
	(2.51)	(2.76)
Market value	0.0006	0.0008
	(0.36)	(0.48)
Home illiquidity	-0.0019	-0.0014
	(-1.09)	(-0.79)
US illiquidity	-0.0028**	-0.0033**
	(-2.14)	(-2.47)
Home-market volatility	1.4977***	1.4303***
	(2.75)	(2.64)
US-market volatility	-1.4192***	-1.3324***
	(-2.80)	(-2.66)
Currency volatility	-0.0306	-0.0499
	(-0.23)	(-0.39)
US institutional ownership	-0.0086	-0.0087
	(-1.17)	(-1.23)
Home number of analysts	-0.0005	-0.0005
	(-1.13)	(-1.20)
Home dispersion of analysts	-0.0001	0.0001
	(-0.04)	(0.06)
Constant	0.0476*	0.0425
	(1.81)	(1.63)
Country fixed-effects	Yes	Yes
Quarter fixed-effects	Yes	Yes
Observations	9,654	9,913
R-squared	0.2183	0.2235

2.6.3 Subsample analyses

2.6.3.1 Excluding Canadian and Asia-pacific cross-listed stocks

Table 2.1 shows that Canadian stocks represent the largest group of stocks listed in the US from a single country, which account for 46.07% of the sample. Canadian cross-listed stocks might be different from other cross-listings for several reasons. First, the trading time of the Canadian stock market coincides with the US trading time. Second, Canadian stocks are listed in US market as ordinary shares, while stocks from other countries are usually listed as ADRs. As explained in Pulatkonak and Sofianos (1999), Canadian ordinaries trading in the US are more fungible (or exchangeable) with the home-market shares than the ADRs. To ensure our main finding is not driven solely by the Canadian cross-listings, we reproduce Table 2.5 using non-Canadian cross-listed pairs. The results are reported in Panel A of Table 2.10.

In our sample, Asia-pacific markets have completely non-overlapping trading hours with the US market. Our measures on attention comovement and price deviations are therefore subject to this time-difference limitation. To ensure the results are not driven by these completely non-synchronous pairs, we restrict our sample to stock pairs domiciled outside of the Asia-pacific region. Regression results are reported in Panel B of Table 2.10.

Panel A shows that coefficients on *AttentionComove* are negative and statistically significant in different specifications of Equation (2.6) in non-Canadian cross-listed pairs. Therefore, our finding on the negative relation between attention comovement and price deviation is not driven solely by the Canadian cross-listed pairs. Similarly, Panel B shows that the negative relation between attention comovement and price deviations remains robust after excluding the Asia-pacific cross-listed pairs. This alleviates the concern that our results are biased by the large time-difference between Asia-pacific and US markets.

Table 2.10 Attention comovement and price deviation in subsamples

The table reports the results from different specifications of Equation (2.6) in two subsamples. Panel A reports the results for the non-Canadian cross-listed pairs. Panel B reports the results for the cross-listed pairs domiciled outside of the Asia-pacific region. The dependent variable is the absolute price deviation in a cross-listed pair, measured by averaging $Ln(P_{US}/P_H)_t$ of each day across a quarter and taking its absolute value. AttentionComove is the attention comovement measure specified in Equation (2.4). All control variables are defined in Equation (2.6). We report the panel regression results across all firms and quarters with country and quarter fixed-effects. T-statistics are based on firm-clustered standard errors. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Panel A: Non-Canadian cross-listed pairs					
AttentionComove	-0.0042***	-0.0040***	-0.0040**	-0.0045***	-0.0043***
	(-3.07)	(-2.78)	(-2.55)	(-2.83)	(-2.81)
Idiosyncratic risk		0.8687**	0.8581**	0.7884**	0.8751**
		(2.04)	(2.33)	(2.12)	(2.17)
Dividend yield		-0.1160	-0.0986	-0.0900	0.0578
		(-0.76)	(-0.76)	(-0.70)	(0.39)
Interest rate		0.1901***	0.1863**	0.1510**	0.1563**
		(2.67)	(2.55)	(2.13)	(2.44)
Market value			-0.0031	-0.0031	-0.0005
			(-0.91)	(-0.91)	(-0.17)
Home illiquidity			-0.0079	-0.0077	-0.0037
			(-1.39)	(-1.36)	(-0.46)
US illiquidity			-0.0018	-0.0013	-0.0104**
			(-0.32)	(-0.24)	(-2.40)
Home-market volatility				1.9002***	1.7365***
				(3.88)	(2.99)
US-market volatility				-1.5223***	-1.6052***
				(-3.29)	(-2.94)
Currency volatility				-0.0659	-0.0194
				(-0.60)	(-0.13)
US institutional ownership					-0.0380***
					(-2.67)
Home number of analysts					-0.0006
					(-1.15)
Home dispersion of analysts					0.0001
					(0.03)
Constant	0.2153***	0.0351	0.0595	0.0505	0.0492
	(3.97)	(1.17)	(1.52)	(1.29)	(1.41)
Country fixed-effects	Yes	Yes	Yes	Yes	Yes
Quarter fixed-effects	Yes	Yes	Yes	Yes	Yes
Observations	11,782	9,819	9,817	9,817	6,099
R-squared	0.1282	0.1626	0.1665	0.1742	0.2117

Table 2.10 continued

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Panel B: Non Asia-pacific cross-listed pairs					
AttentionComove	-0.0025***	-0.0012*	-0.0016**	-0.0017**	-0.0011**
	(-3.22)	(-1.89)	(-2.35)	(-2.47)	(-2.00)
Idiosyncratic risk		0.5699***	0.6692***	0.6567***	0.7383*
		(3.68)	(3.22)	(3.13)	(1.95)
Dividend yield		-0.0975	-0.0910	-0.0891	0.0249
		(-0.86)	(-0.90)	(-0.88)	(0.18)
Interest rate		0.2079***	0.2049***	0.1829***	0.1767***
		(3.26)	(3.21)	(2.99)	(3.02)
Market value			-0.0004	-0.0004	0.0028**
			(-0.34)	(-0.37)	(2.00)
Home illiquidity			0.0013	0.0013	0.0002
			(0.53)	(0.53)	(0.14)
US illiquidity			-0.0042*	-0.0043*	-0.0020
			(-1.86)	(-1.87)	(-1.48)
Home-market volatility				1.1690***	1.3552***
				(2.89)	(3.01)
US-market volatility				-1.0352***	-1.3673***
				(-2.86)	(-3.52)
Currency volatility				-0.0783	-0.0298
				(-0.79)	(-0.21)
US institutional ownership					-0.0104*
					(-1.96)
Home number of analysts					-0.0006**
					(-2.30)
Home dispersion of analysts					-0.0004
					(-0.31)
Constant	0.1305***	0.0382*	0.0391*	0.0340	0.0233
	(5.26)	(1.92)	(1.66)	(1.40)	(1.02)
Country fixed-effects	Yes	Yes	Yes	Yes	Yes
Quarter fixed-effects	Yes	Yes	Yes	Yes	Yes
Observations	19,208	16,667	16,640	16,640	7,902
R-squared	0.1343	0.2068	0.2106	0.2139	0.3069

2.6.3.2 Different sample periods

We also investigate the relation between attention comovement and price deviation in different sample periods. Table 2.1 shows that the increasing pattern in US cross-listings shifts in 2004, and the growth in US cross-listings significantly slows down thereafter. Existing literature attributes this observed trend to the implementation of Sarbanes-Oxley Act (SOX) of 2002. Studies show that SOX discourages firms to cross-list in the US, and leads the lower valued firms to seek cross-listings in markets with less stringent regulation (e.g., Bianconi et al., 2013). The potential changes in the characteristics of US cross-listed firms as a result of SOX may affect the relation between attention comovement and price deviation. Accordingly, we divide our sample into two subperiods (1996-2003 and 2004-2016), and investigate the relation between attention comovement and price deviation, separately.

Results in Panel A of Table 2.11 provide supportive evidence for the negative relation between attention comovement and price deviation over the 1996-2003 period. Coefficients on *AttentionComove* are negative and statistically significant in Models (1) to (4). Similarly, Panel B shows a significantly negative relation between attention comovement and price deviation over the 2004-2016 period. Thus, our findings on the negative effect of attention comovement on price deviation is not confined to a particular sample period.

Table 2.11 Attention comovement and price deviation in subperiods

The table reports the results from different specifications of Equation (2.6) in two subperiods. Panel A reports the results for the subperiod of 1996-2003. Panel B reports the results for the subperiod of 2004-2016. The dependent variable is the absolute price deviation in a cross-listed pair, measured by averaging $Ln(P_{US}/P_H)_t$ of each day across a quarter and taking its absolute value. *AttentionComove* is the attention comovement measure specified in Equation (2.4). All control variables are defined in Equation (2.6). We report the panel regression results across all firms and quarters with country and quarter fixed-effects. T-statistics are based on firm-clustered standard errors. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Panel A: 1996-2003 subperiod					
AttentionComove	-0.0040***	-0.0019*	-0.0021*	-0.0022*	-0.0008
	(-2.70)	(-1.76)	(-1.79)	(-1.93)	(-0.82)
Idiosyncratic risk		0.3617***	0.4010*	0.3764*	0.2448
		(2.72)	(1.91)	(1.78)	(1.15)
Dividend yield		-0.2853	-0.2749	-0.2903	0.0017
		(-1.47)	(-1.55)	(-1.63)	(0.01)
Interest rate		0.0878**	0.0868**	0.0373	0.0474*
		(2.19)	(2.16)	(1.11)	(1.83)
Market value			-0.0010	-0.0010	-0.0011
			(-0.41)	(-0.39)	(-0.39)
Home illiquidity			-0.0009	-0.0008	-0.0011
			(-0.70)	(-0.63)	(-0.56)
US illiquidity			-0.0011	-0.0010	-0.0006
			(-0.79)	(-0.75)	(-0.56)
Home-market volatility				1.0485*	1.7666**
				(1.75)	(2.22)
US-market volatility				-0.6929	-1.4685**
				(-1.57)	(-2.44)
Currency volatility				0.2943**	0.2652
				(1.98)	(1.48)
US institutional ownership					-0.0250*
					(-1.96)
Home number of analysts					-0.0002
					(-0.79)
Home dispersion of analysts					-0.0023
					(-1.23)
Constant	0.0548**	-0.0137*	-0.0074	-0.0107	-0.0081
	(2.58)	(-1.69)	(-0.35)	(-0.50)	(-0.40)
Country fixed-effects	Yes	Yes	Yes	Yes	Yes
Quarter fixed-effects	Yes	Yes	Yes	Yes	Yes
Observations	5,119	4,460	4,443	4,443	1,862
R-squared	0.1609	0.2546	0.2552	0.2596	0.3999

Table 2.11 continued

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Panel B: 2004-2016 subperiod					
AttentionComove	-0.0021***	-0.0019**	-0.0020***	-0.0022***	-0.0026**
	(-2.77)	(-2.40)	(-2.60)	(-2.83)	(-2.47)
Idiosyncratic risk		0.6207***	0.5663**	0.5577**	0.7875**
		(2.78)	(2.34)	(2.29)	(2.15)
Dividend yield		-0.0375	-0.0083	-0.0010	0.1292
		(-0.31)	(-0.08)	(-0.01)	(0.92)
Interest rate		0.3791***	0.3782***	0.3504***	0.1740**
		(4.67)	(4.65)	(4.67)	(2.11)
Market value			-0.0011	-0.0011	0.0004
			(-0.82)	(-0.84)	(0.22)
Home illiquidity			0.0023	0.0022	-0.0010
			(0.60)	(0.59)	(-0.43)
US illiquidity			-0.0037	-0.0038	-0.0051***
			(-1.29)	(-1.31)	(-2.63)
Home-market volatility				1.6803***	1.5751***
				(3.88)	(2.82)
US-market volatility				-1.6179***	-1.6700***
				(-3.75)	(-3.10)
Currency volatility				0.0584	0.1530
				(0.76)	(1.35)
US institutional ownership					-0.0054
					(-0.65)
Home number of analysts					-0.0002
					(-0.48)
Home dispersion of analysts					0.0015
					(0.68)
Constant	0.1275***	0.0522***	0.0604***	0.0477**	0.1007***
	(9.78)	(2.84)	(2.63)	(1.98)	(3.80)
Country fixed-effects	Yes	Yes	Yes	Yes	Yes
Quarter fixed-effects	Yes	Yes	Yes	Yes	Yes
Observations	17,073	14,836	14,826	14,826	7,876
R-squared	0.1797	0.2194	0.2223	0.2277	0.2585

2.7 Conclusion

In this chapter, we investigate whether attention comovement exists between the home-market shares and their cross-listed counterparts. Empirical findings reveal the existence of strong attention comovement in these cross-listed pairs. We then examine the determinants of attention comovement. Our results show that the information environment, information shocks, and aggregate market attention play a significant role in driving attention comovement between the home-US stock pairs. Finally, we investigate the capital market implications of attention comovement. We find that attention comovement helps explain deviations from price parity for the cross-listed stocks.

Our study sits at the intersection of two literatures. The literature on investor attention has extensively investigated the effect of limited investor attention on market efficiency and asset pricing, yet little is known about the drivers of investor attention. Our study contributes to the attention literature by providing evidence of alternative explanations for investor attention. Literature on price deviations for cross-listed stocks identify a number of market frictions related factors that can lead to mispricing. However, the information flows channel has not been explored for price deviations. Our study investigates the source of price deviations from information flows channel. In a broader sense, our finding shed some lights on how attention comovement affects price discrepancies between assets that are fundamentally linked.

Chapter 3

News spillover and return comovement

3.1 Introduction

Comovement in asset returns has long been the subject of academic study, and existing studies have uncovered numerous patterns of comovement in asset returns. Theories under the assumptions of no frictions and rational investors suggest that comovement in prices reflect comovement in fundamentals. However, in the presence of frictions or behavioural biases, comovement in prices can be delinked from comovement in fundamentals. A group of studies contend that the observed stock return comovement is too high relative to fundamentals (Pindyck and Rotemberg, 1993; Barberis et al., 2005). Other evidence favours the friction- and sentiment-based explanations on return comovement (Barberis and Shleifer, 2003; Kumar and Lee, 2006; Vijh, 1994).

Roll (1988) posits that the extent to which stocks move together depends on the amounts of firm-level and market-level information being capitalized into stock prices. He concludes that low return comovement measured by R^2 from common asset pricing models reflects "either private information or else occasional frenzy unrelated to concrete information" (p. 56). Consistent with Roll (1988), the extant literature offers contradictory views on the implications of return comovement. The pioneering study of Morck, Yeung, and Yu (2000) shows that stock prices move together more in poor economies than in rich economies. Subsequent to Morck et al. (2000), numerous studies interpret lower return comovement as an indication that more firm-specific information is being incorporated into stock prices (Durnev, Morck, and Yeung, 2004; Durnev, Morck, Yeung, and Zarowin, 2003; Hutton, Marcus, and Tehranian, 2009; Jin and Myers, 2006). This leads to informative prices reflecting firm-specific information. However, a number of studies question this interpretation and conclude the opposite. That is, lower return comovement is associated with higher price uncertainty, suggesting a positive relation between return comovement and price informativeness (Bartram, Brow, and Stulz, 2012; Dasgupta, Gan, and Gao, 2010; Griffin, Kelly, and Nardari, 2010; Hou, Peng, and Xiong, 2013).

The source of comovement in asset prices remains an open question in financial economics. It has important implications for understanding price formation, asset allocation and risk management. In this chapter, we investigate intra-industry information spillover and its implications for return comovement. Our study is built on Veldkamp's (2006) theoretical model of competitive information market. In an information market, suppliers must provide the highest-value signals to be competitive. Since signals that predict many assets' values generate more expected profits for investors, market forces induce suppliers to sell these types of signals to many market participants. When investors price assets using a common subset of information, news of one asset can affect prices of other assets, generating comovement in asset prices. Veldkamp (2006) develops a framework where investors use information from a subset of assets to value other assets when making investment decisions. The framework is consistent with empirical findings that investors use information of industry leaders to learn about other firms in the same industry (Engelberg, Ozoguz, and Wang, 2018; Hameed, Morck, Shen, and Yeung, 2015). Financial practitioners refer to those industry leaders as 'industry bellwether stocks'. [25]

Building on the theoretical prediction of Veldkamp (2006) and motivated by the industry practice, we conjecture that news from bellwether firms is relevant to their industry peers. Furthermore, if information producers (e.g., analysts or institutional investors) optimize information gathering costs by processing more information about bellwether firms and then applying the information to their related peers, we expect bellwether firms' news to be largely incorporated into their industry peers' valuations. In addition, if a stock's news is important for the pricing of many other stocks, this stock should exhibit stronger return comovement with the market. This is because the information of the stock is capitalized into the prices of many stocks, which advocates the price efficiency view of return comovement. We test these predictions in this chapter.

Our study utilizes news data from Thomson Reuters News Analytics (TRNA), which is available for the period January 2003 to March 2016. TRNA collects and analyses firm-level news content from major news outlets such as Dow Jones Newswires, the Wall Street Journal, Reuters, and local newspapers. It not only provides the flow of news articles (i.e., news coverage) related to a firm, but also produces three quantitative scores that represent the probabilities of a news story being positive, negative, and neutral based on textual analysis. TRNA also provides a score of how relevant a news item is to a firm. If there is a mention of multiple firms in a news article, the firm with the most mentions will have the highest relevance score. We quantify firm news using both news coverage and news tone score, where news tone score is calculated as the difference between positive and negative scores and weighted by the news' relevance score to a firm (Hendershott, Livdan, and Schurhoff, 2015).

[25] According to Investopedia.com, a bellwether stock is a stock believed to be a leading indicator of the direction of the economy, a sector of the market or the market as a whole. For example, analysts use information about Intel to infer the performance of other firms in the technology sector.

Following Hameed et al. (2015), we define industry bellwether stocks as those attracting the most analyst following and whose fundamentals best predict those of industry peer firms. Using a comprehensive dataset of 9,537,249 news stories, we find strong evidence of news spillover from industry bellwether firms to their industry peers. News about bellwether stocks has significant influences on industry peers' stock prices, trading activity and analyst forecasts. Specifically, stock returns and analyst earnings forecast revisions on industry peers are positively associated with bellwether firms' news tone. Trading volume and analyst forecast accuracy of industry peers are positively related to bellwether firms' news coverage. Importantly, this information spillover is unidirectional, that is, news about non-bellwether firms has no effect on bellwether firms. The observed information spillover from bellwether firms to their industry peers is consistent with Veldkamp's (2006) prediction that investors use a common subset of information to predict other assets' values.

Motivated by the prominent role of bellwether firms' news, we next explore why news about bellwether firms is important. To address this question, we investigate the relation between news of bellwether firms and news of their peer firms. Using partial correlation in news to capture the contribution of a firm's news in explaining its industry peers' news, we find that bellwether firms have stronger news partial correlations compared to their industry peers. This finding is in line with the view that information producers cluster their coverage on leading firms. It explains the unidirectional information spillover from the news production channel.

Having established that investors use news of more informative firms to learn about other firms. We then link the degree of a firm's partial correlation in news to price efficiency, proxied by return comovement between the firm and the market (Jin and Myers, 2006; Morck et al., 2000; Roll, 1988). Using R^2 obtained from regressing individual stock returns on the market, we find that stock return comovement is positively associated with news partial correlation. This finding suggests that intra-industry information production plays an important role in explaining return comovement. Furthermore, since stocks with high news partial correlation are more informative relative to other stocks, our finding supports the view that high return comovement is associated with more informative prices.

To reaffirm the implication of our finding for price efficiency from return comovement, we explicitly examine the relation between news partial correlation and stock mispricing. Mispricing is measured by the mispricing score in Stambaugh, Yu, and Yuan (2015), which is constructed by combining the mispricing elements across 11 well documented anomalies. Empirical result reveals a negative relation between news partial correlation and mispricing score. Thus, our finding supports the positive relation between return comovement and price informativeness.

In summary, this chapter documents unidirectional news spillover from bellwether firms to other firms in the same industry. News about bellwether stocks exhibits significant influence on peer stocks, but not the other way around. Also, we find bellwether firms' news contributes to explaining their industry peers' news. Furthermore, we show that news correlation is positively associated with return comovement and price efficiency.

This study, to the best of our knowledge, is the first to establish the relation between intra-industry information spillover and stock return comovement. Our study makes several contributions to the literature. First, our finding that investors use news of bellwether stocks to value other stocks validates the empirical implication of Veldkamp's (2006) theoretical framework. Second, we contribute to the growing body of literature on the effect of social media in financial markets by showing that news production has a significant impact on stock price formation. Third, our results add to the long-standing debate on the implications of return comovement by providing evidence for the positive relation between return comovement and price efficiency. Veldkamp (2006) notes that without data on investors' information, information-driven comovement cannot be tested directly. Using firm-level news as a direct measure of firm-specific information available to investors, we fill this gap and explicitly test how a stock's informativeness is associated with its return comovement and price efficiency.

The remainder of this chapter is structured as follows. Section 3.2 reviews the literature on stock return comovement and news in financial markets. Section 3.3 summarises our research questions and presents hypotheses. Section 3.4 describes the data sources and methodology. Sections 3.5, 3.6 and 3.7 present the empirical results and robustness checks, respectively. Section 3.8 concludes the chapter.

3.2 Literature review

3.2.1 Comovement in stock returns

There is a vast amount of empirical evidence on comovement in asset returns. For example, there are common factors in returns of small stocks, value stocks, stocks with similar prices, and stocks in the same industry and market index (Barberis and Shleifer, 2003; Green and Hwang, 2009; Jame and Tong, 2014; Vijh, 1994). There is also comovement of individual stocks within national markets and among international markets (Bekaert, Hodrick, and Zhang, 2009; Berben and Jansen, 2005). Existing literature provides two broad theories to explain why assets comove.

Finance theories derived from economies without frictions and with rational investors suggest that comovement in prices reflects comovement in fundamentals. The prices of two assets move together only in response to common shocks to fundamentals. One example is Ross's (1976) arbitrage pricing theory where deviations of prices from fundamentals is limited by the presence of arbitrageurs. Thus, investor demand that is not driven by fundamentals is irrelevant. Despite the importance of these theories in modelling the pricing process, they are difficult to reconcile with the abundant evidence that security prices move together either too little or too much relative to their fundamentals.

Alternative theories argue that, due to market frictions or noise-trader sentiment, return comovement can be delinked from fundamentals. This view is supported by many empirical studies. For example, Pindyck and Rotemberg (1993) show that stock price comovement of companies in unrelated business lines is too high to be justified by the covariance in their fundamentals. Froot and Dabora (1999) study "Siamese twin" companies which have exposures to same firm fundamentals and find that returns on shares of these companies traded on different exchanges were far from being perfectly correlated. Barberis and Shleifer (2003) show that assets in the same style comove too much and assets in different styles comove too little. Reclassifying an asset into a new style raises its correlation with that style. Using the changes in index membership, several studies document a significant difference in return comovement within the stocks in the index, before and after index deletions/additions (Barberis et al., 2005; Greenwood and Sosner, 2007; Greenwood, 2007). These findings support the view that comovement of stock returns can be a consequence of the commonality in trading behaviour or investor sentiment rather than the commonality in fundamentals.

In sum, while the true cause of comovement in asset prices is debatable, empirical evidence shows that it deviates from the covariance of their fundamentals

(Barberis et al., 2005; Pindyck and Rotemberg, 1993). Roll (1988) attributes the low R^2s from asset pricing models to high firm-specific return variation not associated with the release of public information. He therefore concludes that a low R^2 seems "to imply the existence of either private information or else occasional frenzy unrelated to concrete information" (p. 56). This motivates subsequent studies to explore the information implication of stock return comovement. Morck et al. (2000) present empirical evidence that stock return comovement is higher in countries with weaker property-rights protection. The authors argue that poor protection of private property rights impede the capitalization of firm-specific information into stock prices, resulting in a higher degree of stock return comovement.

Complementing the findings of Morck et al. (2000), Jin and Myers (2006) observe that stock return comovement is greater in countries with a more-opaque information environment. Wurgler (2000) shows capital allocation is more efficient in markets with low comovement in stock returns. Bushman, Piotroski, and Smith (2004) show that stock returns exhibit greater firm-specific return variation in countries with more developed financial industries and with more free presses. Campbell, Lettau, Malkiel, and Xu (2001) also find a secular decline in return comovement in the United States from 1960 to 1997. Overall, these findings suggest that low return comovement is associated with a higher degree of overall market efficiency.

Extending the market-level analyses to firm-level analyses, studies document a strong link between comovement in stock returns and price informativeness. Durnev et al. (2003) show that firms and industries exhibiting low return comovement with the market incorporates more information about future earnings in current stock prices. Durnev et al. (2004) document a positive relation between corporate investment efficiency and the magnitude of firm-specific information in stock returns. Chen, Goldstein, and Jiang (2006) show that price non-synchronicity increases the sensitivity of investment to prices, as prices provide more information to managers in their investment decision. Hutton et al. (2009) use earnings management as a measure of opacity and find that opacity is associated with a higher degree of return comovement. These firm-level findings further support the view that higher firm-specific return variation signals more information-laden stock prices and, therefore, more efficient stock markets.

Contrary to the aforementioned view that low return comovement proxies for high price efficiency, the literature on costly arbitrage suggests that, due to limits to arbitrage, low stock return comovement is associated with a higher degree of mispricing. Pontiff (2006) argues that idiosyncratic volatility is a form of arbitrage cost. High exposure to idiosyncratic risk forces arbitrageurs to take limited positions in mispriced securities, which impedes market efficiency. Consistent with this costly-arbitrage explanation, Kelly (2014) and Teoh, Yang, and Zhang (2009) show that high idiosyncratic volatility is associated with a poor firm-level information environment. Mashruwala, Rajgopal, and Shevlin (2006) also document that anomalies are higher for stocks with high firm-specific return variation, and argue that idiosyncratic risk imposes barriers to exploiting mispricing.

Apart from the view on costly arbitrage, most recent studies on return comovement and price informativeness also cast doubt on the argument that low return comovement is a proxy for price efficiency. Dasgupta et al. (2010) argue that high return synchronicity might be associated with more informative stock prices, as there is little surprise when the events actually happen in the future. Chen, Huang, and Jha (2012) and Rajgopal and Venkatachalam (2011) provide direct support for Dasgupta et al. (2010) in that high return comovement is associated with better information quality, using different measures of firm-level information quality. Chan and Chan (2014) show that higher stock return synchronicity reflects a better information environment at the time of the SEO. Chan, Hameed, and Kang (2013) document a positive relation between stock return synchronicity and liquidity, arguing that stocks with less firm-specific information face less information asymmetry. Hou et al. (2013) show that low return comovement is associated with more pronounced medium-term price momentum and long-term price reversal, which cautions against using return comovement as a measure of market efficiency.

The mixed findings also exist at the country level. Bartram et al. (2012) show an inconsistent relation between return comovement and country characteristics, which contradicts the positive relation between return comovement and stock market development documented by previous studies. Similarly, Griffin et al. (2010) show that stock return comovement is not associated with a country's institutional quality, as suggested by Morck et al. (2000).

Veldkamp (2006) provides an explanation, within a rational expectation framework, for the positive relation between return comovement and price informativeness observed in empirical studies. She argues that information is a nonrival good, and there is a fixed cost in producing information. For this reason, information producers charge more for low-demand information than for high-demand information. Rationally, the lower price of high-demand information makes the investors want to purchase the information from a common subset of assets which may be useful for predicting the value of many other assets. As a result, news about one asset affects prices of other assets, generating comovement in asset prices. Complementing Veldkamp's (2006) explanation, Mondria (2010) provides a model in which attention constrained investors observe information about a combination of assets rather than individual asset. When investors use this information to predict the value of multiple assets, changes in one asset affect prices of other assets and may lead to asset price comovement.

Hameed et al. (2015) provide the most direct empirical evidence for the theoretical predictions in Veldkamp (2006). They find that investors choose to observe signals that are good predictors of many assets. The authors designate the high analyst coverage firms whose fundamentals best predict their industry peers as bellwether firms. They show that information related to bellwether firms is useful in predicting the prices of more opaque stocks. Therefore, bellwether firms exhibit stronger return comovement because they are priced more accurately.

3.2.2 Information transmission across firms

This chapter is also related to information transmission across firms. There has been ample empirical evidence suggesting that investors face sizeable frictions, and information sometimes transmits slowly in the market place. Many studies document a lead-lag effect in equity markets, in which some firms' stock prices show a delayed reaction to price innovations of other firms.

Lo and MacKinlay (1990) first document a lead-lag effect where returns of small firms are correlated with past returns of big firms, but not vice versa. Hou (2007) finds that this lead-lag effect contains a persistent and highly significant industry component. Big firms lead small firms within the same industry, and it explains the overall lead-lag effect. Hou's (2007) finding is consistent with the hypothesis that information diffuses slowly across firms, and that industries are the primary channel for news dissemination in the equity markets. Cohen and Frazzini (2008) study information transition between customer-supplier linked firms. They find that shocks to one firm translate into shocks to the linked firm in both real quantities (i.e., profits) and stock prices. Similarly, Menzly and Ozbas (2010) find that stocks that are in economically related supplier and customer industries cross-predict each other's returns. Cohen and Lou (2012) investigate how the same piece of industry-specific information affects two sets of firms, when one set of firms requires straightforward processing to update prices, while the other set requires more complicated analyses due to diversified business (i.e., conglomerate firms). They document substantial return predictability from the set of easy-to-analyse firms to their more complicated peers.

Another strand of literature on information transmission is based on the view that when information is costly, the amount of information impounded in prices will directly reflect the cost of information and investors' choices on which assets to learn about (Grossman and Stiglitz, 1980; Veldkamp,2006). Hameed et al. (2015) document a strong spillover effect from high-analyst firms to their fundamentally related peer firms in the same industry. In line with Hameed et al. (2015), Box and Shang (2018) show that investors demand information about firms whose payoffs covary strongly with many others. Engelberg et al. (2018) find evidence that information producers use information about one large firm in an industry cluster to learn about other firms.[26] The authors interpret this evidence to be consistent with a manager's choice to optimize the cost of gathering information by first learning about a large firm within an industry cluster and then applying the correlated information to other firms within the same locality. Overall, empirical evidence supports the view that investors use more informative assets to learn the value of other assets.

3.2.3 News in financial markets

News is one direct measure of information and there has been strong evidence of its important role in financial markets. Earlier studies such as Klibanoff, Lamont, and Wizman (1998) show that country-specific news reported on the front page of the *New York Times* affects the pricing of closed-end country funds. Tetlock (2007) is the first to find evidence that news content can predict movements in stock prices and trading activity. Media pessimism predicts downward pressure on market prices and subsequent reversals. Tetlock (2010) also documents that public news negatively predicts return reversals and positively predicts volume-induced return momentum. The findings are consistent with the argument that public news resolves information asymmetry between the informed and uninformed investors and facilitates more absorption of liquidity shocks. Tetlock, Saar-Tsechansky, and Macskassy (2008) find that the words contained in news reports capture aspects of firm fundamentals that are hard to capture with the easily quantifiable traditional measures of firm performance (e.g., analyst forecasts and accounting variables), and that the content of information embedded in news stories is quickly incorporated into stock market prices. Fang and Peress (2009) document a cross-sectional relation between media coverage and security returns. The study finds that media coverage is positively related to idiosyncratic volatility, suggesting that media coverage expedites the incorporation of firm-specific

[26] They find that fund managers who tilt their portfolios toward large firms within a given industry cluster also tend to hold a larger number of smaller-sized firms within the same industry cluster.

information into prices. In an international context, Griffin, Hirschey, and Kelly (2011) provide evidence that news is more useful in explaining idiosyncratic stock return variation in developed markets than in emerging markets.

The aforementioned studies show that news is associated with substantial price responses in the market. Recent studies suggest that the news media itself has the power to influence financial markets. Engelberg and Parsons (2011) attempted to disentangle the impact of news reporting from the events being reported. The study shows that the news story's very existence – a media effect – is likely to drive trade and price movement. Dougal, Engelberg, Garcia, and Parsons (2012) identify a casual relation between media reporting and aggregate stock market performance. They find that short-term returns on the Dow Jones Industry Average (DJIA) can be predicted using the Wall Street Journal's "Abreast of the Market" (AOTM) column. Since individual columnists are unlikely to possess information relative to the market as a whole, the return predictability may suggest financial journalists can amplify and attenuate investor sentiment, affecting stock market performance. Dougal et al. (2012) provide evidence for Shiller (2000) who notes that "the history of speculative bubbles begins roughly with the advent of newspapers" (P.85). It implies that news media can manipulate investor beliefs apart from fundamentals, their actions and incentives play an important role in prices and asset allocations.

3.3 Research questions and statement of hypotheses

3.3.1 Research questions

Motivated by the existing literature, this chapter investigates how information is disseminated from industry bellwether firms to their peers and its implication for stock return comovement. The study aims to answer three research questions: (1) Is there information spillover among firms within the same industry? (2) If there is information spillover, how is information disseminated across firms? (3) To what extent does intra-industry information spillover contribute to stock return comovement?

3.3.2 Hypotheses

In Veldkamp's (2006) theoretical framework, when information is costly, investors use information from a subset of assets to value other assets. Empirical evidence shows that investors use information about large and heavily analyzed firms (e.g., industry leaders) to learn about their industry peers. Hameed et al. (2015) designate highly followed firms whose fundamentals best predict those of their industry peer firms as industry bellwether firms. If bellwether firms contain information that is useful for other stocks in the same industry, news of bellwether firms should be relevant to their industry peers. Accordingly, our first hypothesis is:

H1: There exists a news spillover effect from bellwether firms to their industry peer firms.

Investors' demand for high-information-content signals induces information suppliers to sell these signals to market participants. Studies show that in order to optimize the cost of gathering information, information producers process more information about firms whose payoffs covary strongly with other firms and then apply the correlated information to the related firms. Bellwether firms are heavily followed by analysts, and their fundamentals are more correlated with their peer firms. As such, it is reasonable to expect that, within the same industry, news of bellwether firms contributes to other firms more than their non-bellwether peers. In this chapter, we use a firm's partial correlation in news (with other firms) to capture the firm's contribution in explaining news of other firms. Our second hypothesis is:

H2: Bellwether firms exhibit a higher partial correlation in news.

The theoretical work of Veldkamp (2006) implies that information production affects stock return comovement. In the previous hypotheses, information of a firm is important for other firms when its fundamentals are correlated with other firms. It is plausible these firms are larger in size and exhibit higher analyst coverage, and therefore are more influential to other firms in the market. It is also reasonable to expect that the prices of these firms are relatively more informative. Linking these expectations to recent empirical findings on the positive relation between return comovement and price efficiency, we conjecture that firms exhibit higher return comovement with the market when their news is influential for other firms. Accordingly, our third hypothesis is:

H3: Firms with more contributing news exhibit higher return comovement with the market.

Building on the same argument, we also expect those firms to exhibit a lower degree of mispricing. Our fourth hypothesis is therefore:

H4: Firms with more contributing news exhibit a lower degree of mispricing.

3.4 Data and methodology

3.4.1 Data

Our sample includes all NYSE, Amex, and Nasdaq-listed firms that are at the intersection of Center for Research in Security Prices (CRSP) for the market data, Compustat Quarterly Fundamentals file for earnings information, IBES for the analyst data, and Thomson Reuters News Analytics (TRNA) for the news data over the period 2003-2016.[27] We exclude non-common stocks (those with CRSP share codes other than 10 and 11). To minimize market frictions, such as price discreteness and bid-ask effects associated with penny stocks, we require the average daily stock price in December of the previous year to be above $1. Our final sample includes 5,454 unique firms.

News data come from the Thomson Reuters News Analytics (TRNA), which is available for the period January 2003 – March 2016. TRNA collects and analyses firm-level news contents from major news outlets such as Dow Jones Newswires, the Wall St Journal, Reuters, and local newspapers.[28] To quantify the information content of a news article, Reuters employs textual analysis and produces three quantitative scores representing the probabilities for a news story being positive, negative, and neutral (the three probabilities sum to one). TRNA also provides a score of how relevant a news item is to a given firm. If a story mentions multiple firms in the contents, the firm with the most mentions is given the highest relevance score. To align the story dates with market data, the date associated with each story is set using a cutoff of the NYSE closing time of 4pm Eastern time. Stories appearing after 4 pm are given the following day's date.

3.4.2 Main variables

3.4.2.1 Partial correlation in fundamentals

The first step in the analysis is to identify bellwether firms based on analyst coverage and partial correlation in fundamentals. Following Hameed et al. (2015), we employ a three-step procedure to construct a measure of fundamental correlation for firm *k* in each quarter. We first run a market model of return on assets (ROA) for each firm in

[27] The availability of news data restricts the sample period.

[28] TRNA has increasingly become popular in academic research, as well as in the industry. For example, Hendershott et al. (2015) employ TRNA to examine whether financial institutions can predict the tone of firm-specific news. Heston and Sinha (2014) study the return predictability of TRNA's news sentiment in the U.S. market. Li, Ramesh, Shen and Wu (2015) show that the coverage of TRNA is comprehensive and covers over 92% of Compustat's earnings announcements. These studies provide detailed description of TRNA.

the industry, other than firm k, using quarterly data over a 5-year window. A minimum of 12 nonmissing quarterly observations is required.

$$ROA_{i,q} = a_i + b_i ROA_{M,q} + e_{i,q} \tag{3.1}$$

where $ROA_{i,q}$ is firm i's ROA in quarter q, and $ROA_{M,q}$ is the asset-weighted ROA across firms, excluding firm k, in the same quarter. Thus, the R^2 from Equation (3.1), $R^2_{i,excl.k}$, is the fraction of variation in firm i's ROA explained by the market.

In the second step, we augment Equation (3.1) with firm k's return on assets, $ROA_{k,q}$:

$$ROA_{i,q} = a_i + b_i ROA_{M,q} + c_i ROA_{k,q} + \varepsilon_{i,q} \tag{3.2}$$

This regression's R^2 is denoted $R^2_{i,incl.k}$. For each pair of stocks (k, i) in the same industry, $R^2_{i,incl.k} - R^2_{i,excl.k}$ is the partial contribution of firm k in explaining firm i's fundamentals, controlling for the market-wide commonality in fundamentals. Industry is defined using 48 Fama and French (1997) industry classifications. We then calculate a partial correlation coefficient as the difference in R^2s normalized by the unexplained fraction of variation in Equation (3.1):

$$PCORR\_ROA_{k,i} = ({R^2}_{i,incl.k} - {R^2}_{i,excl.k})/(1 - {R^2}_{i,excl.k}) \tag{3.3}$$

In the third step, we average $PCORR\_ROA_{k,i}$ across all firms i ($i \neq k$) in the industry, denoting this $PCORR\_ROA_k$. $PCORR\_ROA_k$ is an estimate of firm k's overall fundamentals correlation with all other firms within its industry in each quarter. Because $PCORR\_ROA_k$ is bounded by zero and one, we take a logarithmic transformation:

$$LPCORR\_ROA_k = Ln(PCORR\_ROA_k/(1 - PCORR\_ROA_k)\ ) \tag{3.4}$$

We repeat these steps for each firm in each quarter. A higher $LPCORR\_ROA_k$ means firm k's ROA contributes more to explaining variations in the ROA of other firms in the industry, after controlling for market-wide variations.

3.4.2.2 Partial correlation in news

We measure firm news using both news coverage and news tone. Each day, we count the number of news articles for each stock. $NNEWS_{i,t}$ is computed as the log of one plus the number of news articles about stock *i* on day *t*. The tone score of a news story is the difference between positive and negative scores. To allow for the fact that a news article may be more relevant to one firm than the others, we follow Hendershott et al. (2015) and weight the tone score of a news story by its relevance score to a firm. For each news article, we compute its tone score as follows:

$$NewsTone_i = (Positive_i - Negative_i) * Relevance_i \tag{3.5}$$

We then construct a daily measure of news tone score for stock *i* (denoted by $NewsTone_{i,t}$) by averaging the news tone across all news stories for stock *i* on day *t*. To measure the partial correlation in news, we follow the three-step approach as in computing partial correlation in fundamentals. First, we run a market model of news for each firm *i* using daily data over a 3-month window, excluding firm *k* in the market portfolio.

$$News_{i,t} = a_i + b_i News_{M,t} + e_{i,t} \tag{3.6}$$

where $News_{i,t}$ is either of the two news measures, *NNEWS* or *NewsTone*, for firm *i* on day *t*. $News_{M,t}$ is the news measure for the market portfolio, calculated as the equally weighted average of news measure for all firms with available data on day *t*, excluding firm *k*. The resulting R^2 from Equation (3.6) is the fraction of variation in firm *i*'s news explained by the market.

We then include news for firm *k* on day *t* ($News_{kt}$) in the regression:

$$News_{i,t} = a_i + b_i News_{M,t} + c_i News_{k,t} + e_{i,t} \tag{3.7}$$

Following the same method as specified in Equation (3.3), we calculate the partial correlation in news, $PCORR\_NEWS_{k,i}$, between firms *k* and *i* in each month.

Finally, we average $PCORR\_NEWS_{k,i}$ across all firms *i* within the industry and take a log transformation, denoting this $LPCORR\_NEWS_k$. $LPCORR\_NEWS_k$ captures firm *k*'s overall news correlation with all other firms within its industry in each month. The corresponding news correlations for the two news measures are denoted as: $LPCORR\_NNEWS$ and $LPCORR\_TONE$, respectively.

3.4.3 Methodology

3.4.3.1 News spillover

Our tests require the identification of industry bellwether firms. We utilise the method in Hameed et al. (2015) to identify bellwether stocks. Each year, we first sort stocks in each industry into terciles based on the number of analysts underpinning the one-fiscal-year-ahead earnings forecasts.[29] Within the top analyst coverage tercile, we further rank stocks into three equal groups based on their partial correlation in fundamentals ($LPCORR\_ROA$).[30] We designate stocks with high $LPCORR\_ROA$ among those high analyst coverage stocks as industry bellwether firms.[31] On average, we have 7 bellwether stocks in each industry.[32]

For each industry and each month, we calculate the industry bellwether firms' news tone (number of news) by averaging the news tone scores (number of news items) across bellwether firms, denoting it as $NewsTone_{IBW,t}$ ($NNews_{IBW,t}$). We first gauge the influence of $NewsTone_{IBW,t}$ on returns ($R_{k,t}$) and analyst forecast revisions ($FR_{k,t}$) of industry-peer firm k, using panel regressions of the form:

$$\begin{aligned} R_{k,t}\ (or\ FR_{k,t}) = {} & \alpha_0 + \beta_1 NewsTone_{IBW,t} + \beta_2 NewsTone_{k,t} + \beta_3 R_{k,t-1} + \\ & \beta_4 R_{k,t-2,t-7} + \beta_5 Size_{k,t-1} + \beta_6 BM_{k,t-1} + \\ & \beta_7 Turnover_{k,t-1} + \beta_8 IO_{k,t-1} + \beta_9 R_{m,t} + \\ & \beta_{10} R_{m,t-1} + \varepsilon_{k,t} \end{aligned} \tag{3.8a}$$

where k indexes all firms in each industry save the bellwether firms. $R_{k,t}$ is the stock return of firm k in month t. $FR_{k,t}$ is the earnings forecast revision for firm k in month t, calculated as the change in the mean forecast of 1-year ahead earnings per share from month t-1, scaled by firm k's stock price at the end of month t-1. $NewsTone_{IBW,t}$ is the average news tone of industry bellwether stocks (same industry as stock k). We also control for firm k's own news tone in month t, $NewsTone_{k,t}$.

The regression also includes several other control variables, including the logarithm of firm k's market capitalization at the end of month t-1 ($Size_{k,t-1}$); the log of

[29] We use the number of analysts in the consensus forecasts in December of each year. Our results (as reported in Section 3.7.1) are unaffected if we identify bellwether firms on quarterly basis and measure a firm's analyst coverage using the average monthly number of estimates during the quarter.
[30] Since the industry bellwether firms are identified on an annual basis, we calculate the annual partial correlation in fundamentals for each firm by averaging the quarterly $LPCORR\_ROA$ over each year.
[31] To be included in the analysis, we require at least nine firms are covered by analysts in each industry-year.
[32] The number of bellwether stocks differs significantly across industries, and Appendix A3.1 shows the details.

its book-to-market equity ratio ($BM_{k,t-1}$); its average daily share turnover in month t-1 ($Turnover_{k,t-1}$); and the fraction of shares outstanding held by institutional investors ($IO_{k,t-1}$).[33] In addition, we control for the stock k's returns over the past 6 months ($R_{k,t-2,t-7}$) for price momentum, and its lagged return over the previous month ($R_{k,t-1}$) for short-term reversals. Finally, to control for the impact from the aggregate market, we include the monthly CRSP value-weighted market portfolio return in months t ($R_{m,t}$) and *t-1* ($R_{m,t-1}$).

β_1 captures the stock price reaction of industry peers to bellwether firms' news tone. In the main analysis, we focus on the concurrent price response from industry peers. Since bellwether firms have frequent news coverage, we expect the influence of news in each period to be interim.[34] If investors use news on bellwether firms to update the prices of other related firms, β_1 is expected to be significantly positive. Similarly, if analysts use information about bellwether firms to infer changes in the fundamental value of other firms, we expect positive (negative) news on bellwether firms to lead to upward (downward) analyst forecast revisions on other firms in the industry.

Apart from news tone, we also examine the impact of bellwether firms' news coverage on industry peers. We argue that news tone and news coverage capture different dimensions of news. News tone reveals the news sentiment, while the intensity of news coverage reflects the amount of information available from bellwether firms. If investors trade on information about bellwether firms, we expect bellwether firms' news coverage has significant influence on the trading activity of their industry peers. Furthermore, if bellwether firms are useful to predict values of other firms, more available information about bellwether firms should increase the forecast accuracy on peer firms. Thus, we examine the influence of $NNews_{IBW,t}$ on the trading volume shock ($VOSHOCK_{k,t}$) and analyst forecast accuracy ($FA_{k,t}$) of industry peer firm k, using panel regressions of the form:[35]

$$\begin{aligned} VOSHOCK_{k,t}\ (or\ FA_{k,t}) = \alpha_0 &+ \beta_1 NNEWS_{IBW,t} + \beta_2 NNEWS_{k,t} + \beta_3 R_{k,t-1} + \\ &\beta_4 R_{k,t-2,t-7} + \beta_5 Size_{k,t-1} + \beta_6 BM_{k,t-1} + \\ &\beta_7 Turnover_{k,t-1} + \beta_8 IO_{k,t-1} + \beta_9 R_{m,t} + \\ &\beta_{10} R_{m,t-1} + \varepsilon_{k,t} \end{aligned} \tag{3.8b}$$

[33] We follow Fama and French (1992) to measure book-to-market equity ratio and allow a minimum of 6 months gap between the end of the fiscal year and the price date.
[34] We further look at the lead-lag relation between bellwether firm news and industry peers' share prices in Section 3.6.1.
[35] Conventionally, trading activity is captured by trading volume, order imbalance, volatility, etc. We use trading volume shock as the proxy for trading activity, as Barber and Odean (2008) argue that trading volume in the firm's stock is likely to be greater than usual when value-relevant news about a firm reaches many investors.

Following Bali, Peng, Shen, and Tang (2013), $VOSHOCK_{k,t}$ is the trading volume shock for firm k in month t, measured as:

$$VOSHOCK_{k,t} = \frac{VO_{k,t} - AVGVO_{k,t-12,t-1}}{STDVO_{k,t-12,t-1}} \tag{3.9}$$

where $VO_{k,t}$ is the volume traded for stock k in month t divided by the number of shares outstanding; $AVGVO_{k,t-12,t-1}$and $STDVO_{k,t-12,t-1}$ are the mean and standard deviation, respectively, of the volume traded divided by the number of shares outstanding for stock k over the past 12 months.

In the spirit of Lang, Lins, and Miller (2003), analyst forecast accuracy, $FA_{k,t}$, is defined as the negative of the absolute value of the analyst forecast error deflated by stock price:

$$FA_{k,t} = -\left|\frac{Actual\ Earning_{k,t} - Median\ Estimate_{k,t}}{Stock\ Price_{k,t}}\right| \tag{3.10}$$

where $Actual\ Earning_{k,t}$ is firm k's actual earnings per share, $Median\ Estimate_{k,t}$ is the median analyst forecast of 1-year ahead earnings per share for firm k extracted from I/B/E/S on a monthly basis.

We expect that the arrival of bellwethers' news is associated with more stock trading activity in their industry peers. That is, a positive relation between $NNEWS_{IBW,t}$ and $VOSHOCK_{k,t}$. More news coverage on bellwether firms provides analysts with more information from which to infer the performance of other related firms, leading to better forecast accuracy. Thus, a positive relation is also expected between $NNEWS_{IBW,t}$ and $FA_{k,t}$. Overall, our first hypothesis predicts that news on bellwether firms has significant influence on stock prices, trading activity and analyst forecasts of industry peer firms.

3.4.3.2 Bellwether firms and news correlation

The previous section investigates whether bellwether stocks are informative for valuing other stocks. To support this finding, we further examine whether news of bellwether firms contributes more to news of other firms, as measured by the news partial correlation (*LPCORR_NEWS*). The intuition behind this investigation is that if news producers minimize information gathering cost by processing more information about bellwether firms and applying it to other related firms, we expect firm news, to a large extent, reflects the news content of the bellwether firms. Accordingly, we estimate the following panel regression:

$$LPCORR\_NEWS_{k,t} = a_0 + \beta_1 IBW\_Dummy_{k,t} + \beta_2 Size_{k,t-1} + \beta_3 BTM_{k,t-1} + \beta_4 Price_{k,t-1} + \beta_5 Ret_{k,t-1} + \beta_6 Ret_{k,t-2,t-7} + \beta_7 RetStd_{k,t-1} + \beta_8 Liquidity_{k,t-1} + \beta_9 IO_{k,t-1} + \varepsilon_{k,t} \quad (3.11)$$

where $LPCORR\_NEWS_{k,t}$ is firm k's partial correlation in news with its industry peers calculated in month t. $IBW\_Dummy_{k,t}$ is a dummy variable taking a value of one if firm k is identified as the industry bellwether firm, and zero otherwise. Hypothesis 2 predicts that bellwether firms exhibit higher news partial correlation compared to their industry peers. Thus, β_1 is expected to be significantly positive. In addition, we control for a number of firm-specific variables, including firm size (*Size*), book-to-market (*BTM*), stock price (*Price*), past one-month return (*Ret*), cumulative returns over the past 6 months ($Ret_{i,t-2,t-7}$), return standard deviation (*RetStd*), liquidity (*Liquidity*), analyst coverage (*Analyst*) and institutional ownership (*IO*). If information producers focus on more informative stocks as suggested by prior studies (e.g., Fang and Peress, 2009), we also expect news partial correlation to be positively associated with size, liquidity and institutional ownership.

3.4.3.3 News correlation and return comovement

In this section, we directly examine how news correlation is associated with stock return comovement. Following the literature, R^2 from the following market model is the measure of comovement in stock returns:

$$r_{k,t} = \alpha_k + \beta_k r_{M,t} + \varepsilon_{k,t} \quad (3.12)$$

where $r_{k,t}$ is the return of stock k on day t. $r_{M,t}$ is the market return, calculated as the equally weighted average of all individual stock returns on day t, excluding stock k. Equation (3.12) is estimated for each firm in each month using daily returns over the past 3 months, consistent with the estimation of news correlation. Since the value of R^2 is bounded by zero and one, we take the logarithmic transformation of the R^2 measure:

$$RetComove = \ln\left(\frac{R^2}{1 - R^2}\right) \quad (3.13)$$

We investigate the relation between news correlation and stock return comovement by estimating the following regression:

$$RetComove_{k,t} = a + \beta LPCORR\_NEWS_{k,t} + Controls_{k,t-1} + \varepsilon_{i,t} \quad (3.14)$$

where $RetComove_{k,t}$ is firm k's return comovement in month t. $LPCORR\_NEWS_{k,t}$ is the news partial correlation for firm k ($LPCORR\_NNEWS$ or $LPCORR\_TONE$) in month t. Hypothesis 3 predicts a positive relation between return comovement and news partial correlation. Also, we control for firm-specific variables ($Controls$), including size, book-to-market, price, return, return volatility, liquidity, analyst coverage, and institutional ownership.

3.4.3.4 News correlation and mispricing

Section 3.4.3.3 establishes the relation between news partial correlation and return comovement. To test our conjecture that firms with more contributing news are more price informative, we examine the relation between news partial correlation and mispricing, where mispricing is captured by the mispricing score proposed by Stambaugh et al. (2015). Anomalies reflect mispricing and mispricing has common components across stocks. Building on this concept, the authors combine information across 11 well documented anomalies, and construct a factor capturing common element of mispricing. These 11 anomalies include financial distress, O-Score bankruptcy probability, net stock issues, composite equity issues, total accruals, net operating assets, momentum, gross profitability, asset growth, return on assets, and investment-to-assets. At the end of each month, stocks are ranked based on each anomaly, where the highest (lowest) rank is assigned to the anomaly variable associated with the lowest (highest) average return. A stock's mispricing score is the arithmetic average of its ranking percentile for each of the anomalies.[36] It captures the relative mispricing in the cross-section of stocks.

We investigate the relation between news correlation and mispricing by estimating the following regression across firm-months:

$$Misp\_Score_{k,t} = a + \beta LPCORR\_NEWS_{k,t} + Controls_{k,t-1} + \varepsilon_{i,t} \quad (3.15)$$

where $Misp\_Score_{k,t}$ is firm k's mispricing score in month t and $LPCORR\_NEWS_{k,t}$ is the measure of news partial correlation for firm k ($LPCORR\_NNEWS$ or $LPCORR\_TONE$). Hypothesis 4 predicts news partial correlation to be negatively associated with

[36] The data is available on the author's website http://finance.wharton.upenn.edu/~stambaug/

mispricing. We also control for several firm characteristics that are potential determinants of mispricing, denoted as $Controls_{k,t-1}$, including idiosyncratic volatility, size, book-to-market, price, liquidity, analyst coverage and institutional ownership. See Appendix A3.2 for details of these specific variables.

3.4.4 Summary statistics

Table 3.1 reports descriptive statistics for the key variables used in the analysis. Panel A presents the summary statistics for firm characteristics. The average partial correlation in ROA, *PCORR_ROA*, is 10.55%, and varies substantially across firms and over time with a standard deviation of 5.26%. This result is comparable to Hameed et al. (2015) who document a mean *PCORR_ROA* of 11.68% and standard deviation of 5.94% over the period 1984-2011. The statistics for partial correlation in news differ slightly depending on the news measures. *PCORR_NNEWS* has a mean of 11.26%, ranging from a minimum of 3.9% to a maximum of 24.84%, and *PCORR_TONE* has a mean of 10.33%, exhibiting less variations. On average, the return comovement is -2.2136. An average R^2 of 20.97% suggests that more than 20% variation in the individual stock returns can be explained by the market.

The mean value of news coverage is 15 per month. However, it exhibits a significant variation in the cross-section of firms, with a standard deviation of 25 and ranging from a minimum of 1 news article per month to a maximum of 195 news articles per month. In general, there is more positive news than negative news. The average monthly news tone score is 0.1669. The mean market capitalization is $3.83 billion, and the average book-to-market ratio is 0.7346. A typical firm in our sample is covered by about six analysts. Finally, the mean value of the institutional ownership is 32.07%.

Panel B displays the cross-sectional correlations among the key variables. Since size, book-to-market, price, and institutional ownership are logged in the regressions to follow, the correlations are also based on natural logs of these variables. Overall, the correlations among the key variables are low. The two alternative measures of partial correlation in news, *PCORR_NNEWS* and *PCORR_TONE,* have a correlation just above 0.20, suggesting that they do capture different dimensions of news. The correlation is approximately 0.02 between *PCORR_ROA* and *PCORR_NNEWS,* and 0.01 between *PCORR_ROA* and *PCORR_TONE.* The weak correlation between partial correlation in news and partial correlation in fundamentals implies that the intra-industry news transmission cannot be fully explained by correlated fundamentals. In addition, size is positively correlated with share price (0.70) and number of analysts (0.70). Return volatility is negatively correlated with market value (-0.52) and share price (-0.50), and positively correlated with turnover (0.53).

Table 3.1 Summary statistics

This table presents descriptive statistics for the key variables used in the study. Panel A reports summary statistics for firm characteristics. PCORR_ROA measures the partial correlation in ROA as specified in Equation (3.3). PCORR_NNEWS and PCORR_TONE measure a firm's partial correlation in news with other firms in the same industry based on the number of news and the tone of news, respectively. LPCORR_ROA, LPCORR_NNEWS, and LPCORR_TONE are the logarithmic transformations of PCORR_NNEWS, PCORR_NNEWS, and PCORR_TONE, respectively. Return R^2 is the R^2 estimated from the firm's daily stock returns regressed on the equal-weighted market returns over a 3-month period (excluding the firm). RetComove is the return comovement computed by taking the logarithmic transformation of return R^2 ($\ln(R^2/(1-R^2))$). NNews is the number of news articles related to the firm in a month. NewsTone is the firm's average news tone score in a month. Market value is the market capitalization expressed in billions of dollars. Price is the monthly share price. Book-to-market is the book-to-market ratio measured in June of each year. Turnover is average daily share turnover over a month. Liquidity is measured by Amihud (2002) illiquidity ratio using daily stock returns for each month. Return volatility is the monthly standard deviation of daily stock returns. Number of analysts is the number of estimates underpinning the one-fiscal-year-ahead (FY1) earnings forecasts published in I/B/E/S. Institutional ownership is the percentage of a firm's share outstanding held by institutional investors. Panel B reports correlations among the key variables. Correlations are estimated in the cross section each month and then averaged over time. All variables are winsorized at the 1 and 99 percentiles each month before estimating summary statistics and correlations.

Panel A: Firm characteristics

	Mean	Std. Dev.	Min	P25	P50	P75	Max
PCORR_ROA	0.1055	0.0526	0.0337	0.0702	0.0937	0.1270	0.3036
PCORR_NNEWS	0.1126	0.0431	0.0390	0.0829	0.1054	0.1346	0.2484
PCORR_TONE	0.1033	0.0379	0.0376	0.0777	0.0970	0.1220	0.2230
LPCORR_ROA	-2.2400	0.5124	-3.3565	-2.5837	-2.2698	-1.9275	-0.8301
LPCORR_NNEWS	-2.1330	0.4269	-3.2055	-2.4039	-2.1381	-1.8609	-1.1071
LPCORR_TONE	-2.2245	0.4024	-3.2428	-2.4735	-2.2314	-1.9738	-1.2480
Return R^2	0.2097	0.1906	0.0001	0.0379	0.1649	0.3377	0.7241
RetComove	-2.2136	2.2455	-9.8068	-3.2354	-1.6223	-0.6735	0.9651
NNEWS	15.0000	25.0000	1.0000	3.0000	7.0000	16.0000	165.0000
NewsTone	0.1669	0.2869	-0.7635	-0.0023	0.1516	0.3542	0.8131
Market value ($ billions)	3.8325	17.7052	0.0063	0.0968	0.3843	1.6587	65.7518
Book-to-market	0.7346	0.8205	-0.3218	0.3037	0.5521	0.9037	4.2370
Price	22.6729	23.4974	0.4800	6.0600	15.4500	30.9300	130.3100
Turnover	0.0077	0.0088	0.0001	0.0021	0.0051	0.0099	0.0468
Liquidity	1.8493	8.5565	0.0000	0.0011	0.0097	0.1449	68.5321
Return volatility	0.0294	0.0221	0.0060	0.0154	0.0233	0.0360	0.1124
Number of analysts	6.0000	7.0000	0.0000	1.0000	4.0000	8.0000	28.0000
Institutional ownership	0.3207	0.3348	0.0000	0.0114	0.1888	0.6057	1.0000

Table 3.1 continued

Panel B: Correlation matrix

		1	2	3	4	5	6	7	8	9	10	11	12
1	PCORR_ROA	1											
2	PCORR_NNEWS	0.0170	1										
3	PCORR_TONE	0.0141	0.2097	1									
4	RetComove	0.0236	-0.0048	-0.0382	1								
5	Market value	0.0176	-0.1880	-0.2429	0.2984	1							
6	Book-to-market	0.0709	-0.0309	-0.0353	0.0660	-0.0263	1						
7	Price	0.0420	-0.0668	-0.1040	0.2432	0.7003	-0.0726	1					
8	Return volatility	0.0168	0.0375	0.0839	-0.2661	-0.5196	-0.0703	-0.5024	1				
9	Turnover	0.0407	-0.0100	-0.0003	-0.1349	-0.1397	-0.0961	-0.1290	0.5292	1			
10	Liquidity	0.0009	0.0176	0.0417	-0.2109	-0.2554	0.0109	-0.2041	0.1440	-0.0913	1		
11	Number of analysts	-0.1685	-0.1179	0.0587	0.1806	0.6970	-0.1020	0.4171	-0.2722	0.0849	-0.1615	1	
12	Institutional ownership	-0.0120	0.0143	-0.0267	0.2384	0.2781	0.0031	0.3333	-0.2308	0.0685	-0.2271	0.2529	1

3.5 Empirical results

3.5.1 Intra-industry news spillover

Hypothesis 1 predicts news spillover from bellwether firms to other firms in the same industry. As described in Section 3.4.3.1, bellwether firms are defined as those with a high analyst following and whose fundamentals are most reflective of other firms in the industry. Table 3.2 summarizes the characteristics of bellwether firms. By construction, bellwether firms have a higher value in PCORR_ROA and the most analysts following. On average, *PCORR_ROA* and analyst coverage for bellwether firms are 15.36% and 15, respectively, compared to 10.55% and 6 reported in Table 3.1 for the entire sample. Unsurprisingly, bellwether firms are also larger, with higher price and institutional ownership, more actively traded, and less volatile.

Table 3.2 Summary statistics for bellwether firms

We present the summary statistics of key firm-specific variables for the industry bellwether firms. PCORR_ROA measures the partial correlation in ROA as specified in Equations (3.3). LPCORR_ROA is the logarithmic transformation of PCORR_ROA. Number of analysts is the number of estimates underpinning the one-fiscal-year-ahead (FY1) earnings forecasts. Market value is the market capitalization expressed in billions of dollars. Price is the monthly share price. Turnover is average daily share turnover over a month. Return volatility is the monthly standard deviation of daily stock returns. Institutional ownership is the percentage of a firm's share outstanding held by institutional investors.

	Mean	Std. Dev.	Min	P25	P50	P75	Max
PCORR_ROA	0.1536	0.0533	0.0746	0.1171	0.1424	0.1777	0.3364
LPCORR_ROA	-1.7573	0.3826	-2.5186	-2.0201	-1.7957	-1.5321	-0.6793
Number of analysts	15.000	7.0000	4.0000	10.0000	14.0000	19.0000	36.0000
Market value ($ billions)	12.0962	31.6571	0.1598	1.1797	3.2766	10.1776	172.8915
Price	42.5286	46.3656	3.2300	18.4800	32.3600	52.7200	207.4500
Turnover	0.0118	0.0091	0.0019	0.0058	0.0091	0.0147	0.0473
Return volatility	0.0226	0.0152	0.0062	0.0130	0.0186	0.0273	0.0812
Institutional ownership	0.7638	0.2121	0.0137	0.6646	0.8077	0.9223	1.0000

To test the first hypothesis, we investigate how news of bellwether firms affects the prices, trading activity and analyst forecasts of their industry peers. We run Equations (3.8a) and (3.8b) across all firms and for all months over the sample period from January 2003 to March 2016. Table 3.3 reports the regression results and the standard errors to calculate t-statistics are clustered by industry.[37] We first regress industry peers' returns on bellwether firms' news tone, $NewsTone_{IBW}$. Consistent with our hypothesis, industry peers' stock returns are significantly associated with

[37] We cluster standard errors by industry because the key variables of interest, $NewsTone_{IBW,t}$ and $NNEWS_{IBW,t}$, are measured within each industry.

$NewsTone_{IBW}$ (coefficient= 0.0265, t = 2.83). In line with prior work, most of the control variables in regression (1) are also significantly related to stock returns.

In regression (2), we examine the relation between news of bellwether firms and analyst forecast revisions of their industry peers. $NewsTone_{IBW}$ is positive and significant at the 1% level (t = 4.04). This means analysts tend to make upward (downward) earnings revisions on other firms in the industry when bellwether firms have positive (negative) news. In addition, we document significant positive coefficients on $R_{k,t-1}$ and $R_{k,t-2,t-7}$, suggesting that a stock's short- and median-term performances have predictive power for its analyst forecast revisions.

In regression (3), we investigate whether news of bellwether firms affects trading activity of other firms. It has been well established in the attention literature that investors trade in response of the arrival of news (e.g., Barber and Odean, 2008; Fang and Peress, 2009). If bellwether firms' news affects investors' trading decisions on other stocks in the industry, intense news coverage on bellwether firms should lead to high trading volume of their industry peers. Consistent with this prediction, news of bellwether firms is positively related to trading volume shocks of peer firms (coefficient = 0.0516, t = 3.24). In regression (4), we examine the relation between bellwether firms' news coverage and analyst forecast accuracy of industry peers. Since analysts typically use information about bellwether stocks to infer fundamental values of other related stocks, more news on bellwether stocks should provide analysts with more information for valuations, leading to more accurate forecasts. The documented positive relation between $NNews_{IBW,t}$ and $FA_{k,t}$ provides supportive evidence for this conjecture.

The variations in returns, trading activity and analyst forecasts can be related to firms' own news announcements. In Panel B, we take firm-specific news into account in the analysis. As can be seen, firm-specific news generally carries the same sign as news of bellwether firms. The exception is in regression (4), where we document a negative relation between a firm's analyst forecast accuracy and its own news coverage. One potential reason for this is that, news of non-bellwether industry peers contains less informative contents, adding noises to analyst forecasts. Overall, Panel B suggests that the influence of the bellwether firms' news remains after accounting for firm-specific news. This result provides evidence for the news spillover effect from bellwether firms to their industry peers. News of bellwether firms contains value-relevant information for other firms in the industry, above and beyond information from those other firms' own news.

To test whether the observed news spillover effect is unidirectional, we perform similar analyses after replacing news of bellwether firms with news of non-bellwether firms. Non-bellwether firms are defined as those with the highest analyst coverage but are in the lowest *LPCORR_ROA* group. We denote the news tone and number of news of non-bellwether firms as $NewsTone_{INBW,t}$ and $NNews_{INBW,t}$, respectively. Panel A of Table 3.4 shows that $NewsTone_{INBW,t}$ predicts returns and forecast revisions of other firms in the industry. However, compared with the results in Table 3.3, the influence is weaker in terms of both magnitude and significance. $NNews_{INBW,t}$ shows no impact on their industry peers' trading activity and analyst forecast accuracy. For comparison, we redefine firms from the lowest analyst tercile and also in the lowest *LPCORR_ROA* group as non-bellwether firms. Panel B shows that news of these firms has no impact on their industry peers. Thus, news spillover effect is unidirectional, only from bellwether firms to other firms.

In summary, the findings in Tables 3.3 and 3.4 support a strong news spillover effect from bellwether firms to other firms in the industry. News tone and news coverage of bellwether firms exert significant influence on stock prices, trading volume and analyst forecasts of industry peer firms. Moreover, this news spillover is unidirectional, as news on non-bellwether firms exhibits no influence on other stocks in the industry.

Table 3.3 News of bellwether firms and its impact on industry peers

This table presents the impact of industry bellwether firm news on industry peers' stock prices, trading activity and analyst forecasts. Panel A shows the results without controlling for industry peers' own news. Regressions (1) and (2) report the impact of the news tone ($NewsTone_{IBW,t}$). Regressions (3) and (4) report the impact of the number of news ($NNews_{IBW,t}$). For all firm *k*, excluding the industry bellwether firms, $R_{k,t}$ is firm *k*'s stock return in month *t*. $FR_{k,t}$ is the revision in consensus forecasts of 1-year ahead earnings per share, $VOSHOCK_{k,t}$ is the trading volume shock, and $FA_{k,t}$ is the analyst forecast accuracy. $NewsTone_{IBW}$ and $NNews_{IBW}$ are the average news tone score and number of news for industry bellwether firms (i.e., same industry as firm *k*), respectively. The firm-specific independent variables include firm *k*'s market capitalization ($Size_k$), book-to-market ratio (BM_k), average daily share turnover ($Turnover_k$), the fraction of shares outstanding held by institutional investors (IO_k), the lagged stock return ($R_{k,t-1}$) and the cumulative return over month *t*-7 to *t*-2 ($R_{k,t-2,t-7}$). *Size, BM, Turnover* and *IO* are natural logged. R_m is the value-weighted returns of all stocks in CRSP. In Panel B, we add the news tone ($NewsTone_{k,t}$) and number of news ($NNews_{k,t}$) for firm *k*. The table reports the regression results across firm-months with industry and month fixed effects. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $R_{k,t}$	(2) $FR_{k,t}$	(3) $VolShock_{k,t}$	(4) $FA_{k,t}$
Panel A: Without firm-specific news				
$NewsTone_{IBW,t}$	0.0265***	0.0021***		
	(2.83)	(4.04)		
$NNews_{IBW,t}$			0.0516***	0.0583***
			(3.24)	(2.61)
$R_{k,t-1}$	-0.0194***	0.0143***	0.4942***	0.5783***
	(-4.43)	(10.53)	(3.88)	(8.38)
$R_{k,t-2,t-7}$	0.0019*	0.0037***	0.2898***	0.3367***
	(1.70)	(6.04)	(8.65)	(8.42)
$Size_{k,t-1}$	-0.0012***	0.0002***	-0.0207***	0.1005***
	(-5.84)	(3.20)	(-6.57)	(8.87)
$BM_{k,t-1}$	0.0016***	-0.0002***	0.0324***	-0.0486*
	(4.16)	(-3.20)	(4.13)	(-1.73)
$Turnover_{k,t-1}$	-0.2437***	-0.0815***	11.8964***	-12.7241***
	(-3.82)	(-5.22)	(10.58)	(-4.13)
$IO_{k,t-1}$	0.0032***	0.0001*	-0.0079**	0.1252***
	(8.17)	(2.02)	(-2.67)	(8.52)
$R_{m,t-1}$	-0.2886	0.0835**	-14.9414***	0.9362
	(-1.03)	(2.58)	(-5.59)	(1.37)
$R_{m,t}$	1.3138***	-0.0959***	11.2781***	-1.3799***
	(6.22)	(-4.72)	(9.28)	(-3.12)
Constant	0.0001	0.0069***	-0.4745***	-0.5336***
	(0.01)	(5.57)	(-6.65)	(-5.14)
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	453,542	280,775	455,320	339,185
R-squared	0.1645	0.0517	0.0599	0.0374
Panel B: With firm-specific news				
$NewsTone_{IBW,t}$	0.0266***	0.0020***		
	(2.72)	(4.00)		
$NewsTone_{k,t}$	0.0354***	0.0015***		
	(15.08)	(9.88)		
$NNews_{IBW,t}$			0.0326**	0.0447**
			(2.27)	(2.29)
$NNews_{k,t}$			0.4038***	-0.0799***
			(17.41)	(-8.66)
$Controls_{k,t-1}$	Yes	Yes	Yes	Yes
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	353,848	239,204	355,568	289,076
R-squared	0.1776	0.0540	0.1057	0.0480

Table 3.4 News of non-bellwether firms and its impact on industry peers

This table presents the impact of non-bellwether firm news on industry peers' stock prices, trading activity and analyst forecasts. In Panel A, for each year and each industry, stocks are grouped into terciles by analyst coverage, within the top coverage tercile, stocks are further sorted into three groups based on partial correlation in fundamentals with other stocks in the industry (*LPCORR_ROA*). Non-bellwether firms are high analyst coverage stocks with low *LPCORR_ROA*. In Panel B, non-bellwether firms are defined as low analyst coverage stocks with low *LPCORR_ROA*. For all firm *k*, excluding the industry non-bellwether firms, $R_{k,t}$ is firm *k*'s stock return in month *t*. $FR_{k,t}$ is the revision in consensus forecasts of 1-year ahead earnings per share, $VOSHOCK_{k.t}$ is the trading volume shock, and $FA_{k.t}$ is the analyst forecast accuracy, defined as the negative of the absolute value of the analyst forecast error. $NewsTone_{INBW}$ and $NNews_{INBW}$ is the news tone score and number of news for industry non-bellwether firms (i.e., same industry as firm *k*), respectively. $Controls$ include the control variables specified in Table 3.3. The table reports the regression results across firm-months with industry and month fixed effects. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1)	(2)	(3)	(4)
	$R_{k,t}$	$FR_{k,t}$	$VOSHOCK_{k,t}$	$FE_{k,t}$
Panel A: News spillover from stocks with high analyst coverage and low *LPCORR_ROA*				
$NewsTone_{INBW,t}$	0.0201**	0.0016*		
	(2.23)	(1.94)		
$NewsTone_{k,t}$	0.0357***	0.0021***		
	(16.25)	(5.52)		
$NNews_{INBW,t}$			0.0310	-0.0393
			(1.39)	(-0.98)
$NNews_{k,t}$			0.3295***	-0.0724***
			(18.29)	(-6.53)
$Controls_{k,t-1}$	Yes	Yes	Yes	Yes
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	355,303	241,399	355,593	289,406
R-squared	0.1790	0.0450	0.1035	0.0370
Panel B: News spillover from stocks with low analyst coverage and low *LPCORR_ROA*				
$NewsTone_{INBW,t}$	0.0024	0.0007		
	(0.80)	(0.87)		
$NewsTone_{k,t}$	0.0366***	0.0020***		
	(15.17)	(5.49)		
$NNews_{INBW,t}$			0.0054	0.0152
			(0.44)	(1.25)
$NNews_{k,t}$			0.3295***	-0.0636***
			(17.77)	(-5.83)
$Controls_{k,t-1}$	Yes	Yes	Yes	Yes
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	359,374	248,098	359,657	296,238
R-squared	0.1813	0.0482	0.1107	0.0394

3.5.2 Bellwether firms and partial correlation in news

In the previous section, we show that news on industry bellwether firms plays an important role in explaining variations in stock prices, trading volume and analyst forecasts of their industry peers. However, the effect is not observed for non-bellwether firms. This finding implies that news of bellwether firms is important and contributes in addition to firm-specific news. This section aims to further confirm this finding.

We use partial correlation in news, $LPCORR\_TONE_{k,t}$ and $LPCORR\_NNEWS_{k,t}$, as described in Section 3.4.2.2, to capture the contribution of firm k in explaining its industry peers' news. We regress news partial correlation on the bellwether firm dummy ($IBW\_Dummy_{k,t}$) and other control variables. $IBW\_Dummy_{k,t}$ equals to one if firm k is identified as industry bellwether firm in the year, and zero otherwise (as specified in Equation (3.11)). Results are presented in Table 3.5.

Coefficients on $IBW\_Dummy_{k,t}$ are significantly positive in both regressions. It suggests bellwether firms have more contributing news than their industry peers. In addition, our results show that firms associated with high news partial correlations also have lower book-to-market, higher stock prices, stronger past returns, lower volatility and a higher level of institutional ownership. Counterintuitively, we document a negative relation between firm size and news correlation. A possible explanation is that, we request a minimum of nine daily observations over a three-month window when constructing news partial correlations. This screening process excludes small firms with insufficient news coverage.[38] As a result, size is not crucial in explaining the variations in news correlations. Results are similar using two alternative measures of news correlation.

In summary, Table 3.5 suggests that bellwether firms are associated with higher news partial correlations. This finding provides evidence that news of bellwether firms is more important than that of their industry peers. Also, it explains the unidirectional news spillover from bellwether firms to other industry firms documented in Section 3.5.1 from the news production channel. Thus, our finding provides further support for Veldkamp's (2006) prediction that, when information is costly, market participants use common signals from more informative assets.

[38] We argue that screening out small firms biases us against finding the results. As demonstrated in Section 3.6.2.4, news spillover is most significant among opaque stocks.

Table 3.5 Bellwether firms and partial correlation in news

This table presents how industry bellwether firms differ from other firms in news partial correlation. $LPCORR\_TONE_{k,t}$ ($LPCORR\_NNEWS_{k,t}$) is the news partial correlation measure estimated based on news tone score (number of news) as described in Section 3.4.2.2. News partial correlation for firm *k* captures the partial contribution of firm *k*'s news in explaining the news of its industry peers. $IBW\_Dummy_{k,t}$ is the industry bellwether firm dummy, taking the value of one if the firm is identified as industry bellwether firm, and zero otherwise. $Size$ is the market capitalization. $Price$ is the monthly share price. BM is the book-to-market equity ratio calculated in June of each year. R_{t-1} is the stock return in month *t-1*, and $R_{t-2,t-7}$ is the cumulative return over the past 6 months with a month lag. $RetStd$ is the standard deviation of the daily returns of the month. $Liquidity$ is the Amihud (2002) illiquidity ratio. IO is the fraction of a firm's share outstanding held by institutional investors. *Size, Price, BM, and IO* are natural logged. The table reports the regression results across firm-months with industry and month fixed effects. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $LPCORR\_TONE_{k,t}$	(2) $LPCORR\_NNEWS_{k,t}$
$IBW\_Dummy_{k,t}$	0.0142**	0.0286***
	(2.26)	(3.15)
$Size_{k,t-1}$	-0.0745***	-0.0666***
	(-7.19)	(-6.60)
$BM_{k,t-1}$	-0.0126***	-0.0056
	(-3.09)	(-1.28)
$Price_{k,t-1}$	0.0343***	0.0327***
	(4.52)	(3.59)
$R_{k,t-1}$	0.0134	0.0342**
	(0.71)	(2.13)
$R_{k,t-2,t-7}$	0.0177***	0.0140*
	(3.06)	(2.03)
$RetStd_{k,t-1}$	-1.3559***	-1.8018***
	(-3.57)	(-4.33)
$Liquidity_{k,t-1}$	-0.0006	-0.0014
	(-0.82)	(-1.00)
$IO_{k,t-1}$	0.0061**	0.0164***
	(2.60)	(4.36)
Constant	-1.9338***	-1.6241***
	(-23.18)	(-18.72)
Industry fixed-effects	Yes	Yes
Month fixed-effects	Yes	Yes
Observations	114,680	114,006
R-squared	0.1080	0.0921

3.5.3 News partial correlation and stock return comovement

Hypothesis 3 predicts a positive relation between news correlation and return comovement. The prediction is based on the view that informative stocks should be priced efficiently, and recent empirical evidence suggests a positive relation between return comovement and price efficiency. To test this hypothesis, we estimate different specifications of Equation (3.14) and report the regression results in Table 3.6.

Regression (1) uses $LPCORR\_TONE$ as the news correlation measure to explain return comovement. Consistent with our hypothesis, $LPCORR\_TONE_{k,t}$ is positive and significant at the 1% level (t = 3.05). In addition, results on the control variables show that return comovement is stronger for large size, high book-to-market and more liquid stocks, and stocks with high analyst coverage and institutional ownership. These findings are consistent with the literature that high return comovement is associated with better information environment (Chan and Chan,2014; Chen et al., 2012; Rajgopal and Venkatachalam, 2014), suggesting a positive relation between return comovement and informative price.

Regression (2) augments regression (1) with a news commonality measure ($TONEComove$) as in Dang et al. (2015). News commonality is measured by the R^2 estimated from regressing an individual firm's news score on the equal-weighted market news score [39] The authors argue that news commonality captures the media's production of firm-specific information, and high news commonality suggets a firm's news events have more market wide information content than firm-specific fundamentals. As a result, greater news comovement results in stronger return comovement. Dang et al.'s (2015) finding tend to favour the view that higher return comovement is caused by less capitalization of firm-specific information.

The intuition behind our news partial correlation measure differs from news commonality used in Dang et al. (2015) in two aspects. First, news commonality captures the extent to which the individual firm's news comoves with the aggregate market news, while our news partial correlation captures the news transfer within the industry. Second, the news commonality gauges the market-wide information content in the firm's news events, while partial correlation in news captures the partial contribution of

[39] Dang et al. (2015) use event sentiment score (ESS) from RavenPack News Analytics in their study, which is similar to our NewsTone measure. The paper uses the R^2 from the following regression as an annual measure of firm-level news commonality: $ESS_{i,t} = a_i + \beta_i ESS_{M,t} + \varepsilon_{i,t}$, where $ESS_{i,t}$ is the ESS value for stock i in week t and $ESS_{M,t}$ is the market ESS value in week t, calculated as the equal-weighted average of weekly ESS values across all firms, excluding firm i. We estimate the news commonality on a monthly frequency based on daily news data in our study to make it consistent with the construction of our news partial correlation variable.

a firm’s news in explaining the news of its industry peers. For these reasons, the two measures provide different explanations for stock return comovement. The news commonality measure implies that when less-firm specific information is incorporated into stock prices, returns exhibit stronger comovement with the market. Our partial correlation measure suggests that news producers use the news on a subset of stocks to infer the value changes of other stocks in the same industry. When a stock’s firm-specific information is incorporated into the prices of many other stocks, this stock exhibits stronger return comovement with the market.

Consistent with the finding in Dang et al. (2015), regression (2) reveals a positive relation between news commonality ($TONEComove_{k,t}$) and return comovement. More importantly, our news partial correlation ($LPCORR\_TONE_{k,t}$) remains significantly positive, suggesting that intra-industry news spillover provides an incremental explanation for stock return comovement.

In regression (3), we repeat the same analysis as in regression (1), but using news partial correlation constructed based on the number of news ($LPCORR\_NNEWS_{k,t}$). The finding on the positive relation between return comovement and news partial correlation remains robust in regression (3) and the coefficients on the control variables differ little from those in regression (1). In regression (4), we estimate Dang el al.’s (2015) news commonality measure using number of news ($NNEWSComove_{k,t}$). Consistent with the results in regression (2), news partial correlation remains robust after accounting for news commonality.

Having established the positive relation between return comovement and contemporaneous news partial correlation, we further investigate whether news correlation has predictive power for future return comovement. We modify the specification in Equation (3.14) to examine the lead-lag relation between stock return comovement and news correlation, controlling for the lagged return comovement. The results are displayed in Panel B. Coefficients on the lagged news partial correlation remain positive and statistically significant at the 1% level in all regressions. In addition, return comovement demonstrates a strong time-series autocorrelation in our sample, coefficients on $RetComove_{k,t-1}$ are above 0.70 in all four regressions. To conserve space and focus on the marginal effects of news correlation, Panel B supresses coefficients of the controls. In the untabulated results, coefficients on the control variables remain similar to those in Panel A. Overall, the findings in Table 3.6 provide evidence that news correlation plays an important role in driving stock return comovement.

Table 3.6 Partial correlation in news and stock return comovement

This table presents results from panel regression of stock return comovement on partial correlation in news. Dependent variable $RetComove_{k,t}$ represents the return comovement for stock k in month t, measured using the R^2 from regression of daily stock returns of stock k on equal-weighted market returns and taking the logarithmic transformation. $LPCORR_{TONE}$ and $LPCORR\_NNEWS$ are news partial correlation measures, based on news tone score and the number of news, respectively. $TONEComove$ and $NNEWSComove$ are the news commonality measures as in Dang et al. (2015), defined as the R^2 obtained from the regression of a firm's news scores on the market's news scores. Panel A reports the contemporaneous relation between return comovement and partial correlation in news and Panel B reports the lead-lag relation. We control for the firm's partial correlation in fundamentals ($LPCORR\_ROA$). Other control variables include market capitalization ($Size$), book-to-market ratio (BM), share price($Price$) , monthly stock return ($R_{k.t-1}$) and the cumulative return over the past 6 months with a month lag ($R_{k,t-2,t-7}$), return standard deviation ($RetStd$), the Amihud (2002) illiquidity ratio ($Liquidity$), the number of estimates making the one-year-ahead earnings forecasts ($Analyst$), and the fraction of shares outstanding held by institutional investors(IO). Size, book-to-market, price, analysts and institutional ownership are natural logged. The table reports the regression results across firm-months with industry and month fixed effects. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $RetComove_{k,t}$	(2) $RetComove_{k,t}$	(3) $RetComove_{k,t}$	(4) $RetComove_{k,t}$
Panel A: Contemporaneous relation between return comovement and news partial correlation				
$LPCORR\_TONE_{k,t}$	0.0846***	0.0818***		
	(3.05)	(2.99)		
$TONEComove_{k,t}$		0.0122***		
		(5.23)		
$LPCORR\_NNEWS_{k,t}$			0.0970***	0.0865***
			(3.56)	(3.22)
$NNEWSComove_{k,t}$				0.0175***
				(8.57)
$LPCORR\_ROA_{k,t}$	0.0389	0.0385	0.0391	0.0379
	(1.59)	(1.57)	(1.59)	(1.52)
$Size_{k.t-1}$	0.1279***	0.1312***	0.1290***	0.1314***
	(5.89)	(6.13)	(5.76)	(5.86)
$BM_{k.t-1}$	0.0978***	0.0986***	0.0972***	0.0977***
	(4.59)	(4.66)	(4.62)	(4.68)
$Price_{k.t-1}$	0.0403	0.0397	0.0386	0.0374
	(0.97)	(0.96)	(0.92)	(0.89)
$R_{k.t-1}$	-0.1488**	-0.1504**	-0.1522**	-0.1516**
	(-2.12)	(-2.14)	(-2.14)	(-2.15)
$R_{k,t-2,t-7}$	0.0511	0.0506	0.0501	0.0502
	(1.31)	(1.29)	(1.25)	(1.27)
$RetStd_{k.t-1}$	-10.2827***	-10.2235***	-10.2635***	-10.2155***
	(-8.68)	(-8.64)	(-8.56)	(-8.55)
$Liquidity_{k,t-1}$	-0.0325***	-0.0324***	-0.0322***	-0.0321***
	(-4.16)	(-4.13)	(-4.20)	(-4.20)
$Analyst_{k,t-1}$	0.0932*	0.0923*	0.0829*	0.0808*
	(1.95)	(1.93)	(1.75)	(1.70)
$IO_{k,t-1}$	0.1292***	0.1289***	0.1265***	0.1260***
	(7.59)	(7.59)	(7.73)	(7.71)
Constant	-2.5015***	-2.4827***	-2.4675***	-2.4283***
	(-7.33)	(-7.26)	(-7.36)	(-7.26)
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	108,745	108,742	108,133	108,131
R-squared	0.3166	0.3170	0.3155	0.3161

Table 3.6 continued

Panel B: Lead-lag relation between return comovement and news partial correlation				
$LPCORR\_TONE_{k,t-1}$	0.0294***	0.0287***		
	(3.41)	(3.40)		
$TONEComove_{k,t-1}$		0.0031**		
		(2.42)		
$LPCORR\_NNEWS_{k,t-1}$			0.0453***	0.0442***
			(4.94)	(4.91)
$NNEWSComove_{k,t-1}$				0.0019
				(1.46)
$LPCORR\_ROA_{k,t-1}$	0.0136	0.0135	0.0131	0.0130
	(1.26)	(1.25)	(1.22)	(1.21)
$RetComove_{k,t-1}$	0.7079***	0.7078***	0.7084***	0.7083***
	(64.84)	(64.94)	(63.86)	(63.67)
$Controls_{k,t-1}$	Yes	Yes	Yes	Yes
Observations	108,433	108,429	107,817	107,814
R-squared	0.6610	0.6611	0.6615	0.6615

3.5.4 News partial correlation and stock price efficiency

In the previous section, our results show that stocks whose news is correlated more with other firms in the same industry are associated with stronger return comovement. We argue that this is because firms with a higher level of news partial correlation is more informative relative to other stocks. Investors and news producers use information on more informative stocks to infer other stocks' value. For this reason, our findings support the view that high return comovement is associated with high price efficiency. Motivated by the long-standing debate on whether stock return comovement reflects the degree of stock price informativeness or noise, we conduct a formal investigation of the relation between news partial correlation and stock price mispricing in this section. Table 3.7 displays the regression results of mispricing scores on news partial correlation.

In regression (1), we regress mispricing score on news partial correlation, measured by $LPCORR\_TONE_{k,t}$, and other variables. The coefficient on $LPCORR\_TONE_{k,t}$ equals -0.0094 and is statistically significant at the 1% level (t= 2.93). This finding provides supportive evidence that high news partial correlation is associated with less mispricing. Also, in line with the prior literature, mispricing is positively related to idiosyncratic volatility, and negatively related to size, price and institutional ownership. In regression (2), we further include Dang et al.' s (2015) news commonality measure ($TONEComove_{k,t}$) in the analysis, and the coefficient is significantly negative (-0.0012). This result is noteworthy, as it suggests an alternative explanation to the positive relation between news commonality and return comovement documented in Dang et al. (2015). That is, firms that exhibit higher news commonality are more accurately priced. In regressions (3) and (4), we repeat the analyses using $LPCORR\_NNEWS_{k,t}$ as the proxy for news partial correlation. Results on the negative relation between news partial correlation and mispricing remain robust.

Overall, Table 3.7 suggests that firms with more contributing news are associated with a higher level of price efficiency. Combining with the finding on the positive relation between news partial correlation and return comovement, our results favors the view that high return comovement is an indication of high price informativeness.

Table 3.7 Partial correlation in news and mispricing

This table presents results from panel regression of mispricing on partial correlation in news. Dependent variable $Misp_Score_{k,t}$represents the mispricing score for firm *k* in month *t*, constructed by Stambaugh et al. (2015). $LPCORR_{TONE}$ and $LPCORR\_NNEWS$ are news partial correlation measures, based on news tone score and the number of news, respectively. $TONEComove$ and $NNEWSComove$ are the news commonality measures proposed in Dang et al. (2015), defined as the R^2 obtained from the regression of a firm's news scores on the market's news scores. We control for the firm's partial correlation in fundamentals ($LPCORR\_ROA$). Other control variables include, idiosyncratic volatility ($IVOL$), market capitalization ($Size$) , book-to-market ratio (BM), share price($Price$), the Amihud (2002) illiquidity ratio ($Liquidity$), the number of estimates making the one-year-ahead earnings forecasts ($Analyst$), and the fraction of shares outstanding held by institutional investors(IO). Size, book-to-market, price, analysts and institutional ownership are natural logged. The table reports the regression results across firm-months with industry and month fixed effects. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $Misp\_Score_{k,t}$	(2) $Misp\_Score_{k,t}$	(3) $Misp\_Score_{k,t}$	(4) $Misp\_Score_{k,t}$
$LPCORR\_TONE_{k.t}$	-0.0094***	-0.0092***		
	(-2.93)	(-2.87)		
$TONEComove_{k,t}$		-0.0012**		
		(-2.41)		
$LPCORR\_NNEWS_{k,t}$			-0.0117***	-0.0117***
			(-3.48)	(-3.04)
$NNEWSComove_{k,t}$				-0.0013**
				(-1.98)
$LPCORR\_ROA_{k,t}$	0.0122**	0.0122**	0.0124*	0.0125*
	(2.09)	(2.09)	(1.89)	(1.91)
$IVOL_{k.t-1}$	3.5217***	3.5147***	3.6956***	3.6937***
	(18.48)	(18.47)	(17.55)	(17.56)
$Size_{k.t-1}$	-0.0578***	-0.0581***	-0.0510***	-0.0513***
	(-10.89)	(-10.95)	(-9.11)	(-9.13)
$BM_{k.t-1}$	-0.0176***	-0.0177***	0.0013	0.0012
	(-3.26)	(-3.28)	(0.23)	(0.21)
$Price_{k.t-1}$	-0.0292***	-0.0291***	-0.0253***	-0.0252***
	(-3.65)	(-3.65)	(-2.66)	(-2.65)
$Liquidity_{k,t-1}$	-0.0005	-0.0005	-0.0004	-0.0005
	(-0.50)	(-0.50)	(-0.38)	(-0.40)
$Analyst_{k,t-1}$	0.0522***	0.0522***	0.0416***	0.0415***
	(6.52)	(6.53)	(4.75)	(4.74)
$IO_{k,t-1}$	-0.0072**	-0.0072*	-0.0107**	-0.0107**
	(-1.96)	(-1.96)	(-2.51)	(-2.51)
Constant	4.4821***	4.4800***	4.2508***	4.2478***
	(100.60)	(100.70)	(78.03)	(77.78)
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	94,198	94,197	93,796	93,795
R-squared	0.3293	0.3294	0.2507	0.2507

3.6 Further tests

3.6.1 Lead-lag news spillover

In Section 3.5.1, we show that news of bellwether firms has significant influence on the contemporaneous stock prices, trading activity and analyst forecasts of their industry peers. In this section, we investigate whether there also exists a lead-lag news spillover effect. We examine whether bellwether firms' news has predictive power for industry peers' returns, trading volume and forecast revisions over the subsequent period. To do so, we bring the dependent variables in Equations (3.8a) and (3.8b) one period forward. Results are presented in Table 3.8.

Findings are consistent with the contemporaneous analysis. Regressions (1) and (2) show that bellwether firms' news tone is positively associated with industry peers' expected returns and analyst forecast revisions over the next month. Regressions (3) and (4) indicate that bellwether firms' news coverage is positively related to the industry peers' trading volume shock and analyst forecast accuracy in the subsequent period. Thus, we provide evidence for both contemporaneous and lead-lag news spillover from bellwether firms to industry peers. More importantly, the finding on the lead-lag news spillover alleviates the concern of reverse causality. That is, whether news tone on bellwether firms changes after market participants observing price changes of all other firms in the industry.

Table 3.8 News of bellwether firms and its impact on industry peers: Lead-lag relations

This table presents the impact of industry bellwether firm news on industry peers' stock returns, trading activity and analyst forecasts in the next period. Regressions (1) and (2) use news tone ($NewsTone$) as the news measure and Regressions (3) and (4) use the number of news ($NNews$) as the news measure. $R_{k,t}$ is firm *k*'s stock return in month *t*; $FR_{k,t}$ is the revision in consensus forecasts of 1-year ahead earnings per share; $VOSHOCK_{k,t}$ is the trading volume shock, and $FA_{k,t}$ is the analyst forecast accuracy, defined as the negative of the absolute value of the analyst forecast error. $NewsTone_{IBW,t-1}$ and $NNews_{IBW,t-1}$ are the news tone score and number of news for industry bellwether firms (i.e., same industry as firm *k*) in month *t-1*, respectively. $Controls_{k,t-1}$ represents control variables specified in Table 3.3, including firm *k*'s market capitalization ($Size_k$), book-to-market ratio (BM_k), average daily share turnover ($Turnover_k$), the fraction of shares outstanding held by institutional investors (IO_k), the lagged stock return ($R_{k,t-1}$), the cumulative return over month t-7 to t-2 ($R_{k,t-2,t-7}$), and value-weighted returns of all stocks in CRSP in month *t*-1 ($R_{m,t-1}$) and month *t* ($R_{m,t}$). The table reports the results from panel regressions with industry and month fixed effects. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $R_{k,t}$	(2) $FR_{k,t}$	(3) $VOSHOCK_{k,t}$	(4) $FE_{k,t}$
$NewsTone_{IBW.t-1}$	0.0084***	0.0029***		
	(2.90)	(2.68)		
$NewsTone_{k,t-1}$	0.0048***	0.0014***		
	(4.97)	(4.79)		
$NNews_{IBW,t-1}$			0.0137**	0.0557**
			(2.13)	(2.21)
$NNews_{k,t-1}$			-0.0240***	-0.0745***
			(-7.49)	(-7.01)
$Controls_{k,t-1}$	Yes	Yes	Yes	Yes
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	356,987	239,162	357,275	290,488
R-squared	0.1355	0.0474	0.0878	0.0381

3.6.2 News spillover under different circumstances

In this section, we further investigate the intra-industry news spillover conditioning on the industry competition, market states, the sign of news tone and industry peers' information environment.

3.6.2.1 News spillover and industry concentration

We first explore the link between intra-industry news spillover and industry concentration. Following the literature (e.g., Hou and Robinson, 2006; Gu, 2016), we measure industry concentration using the Herfindahl-Hirschman Index (HHI), which is defined as the sum of squared market shares:

$$HHI_j = \sum_{k=1}^{K} S_{kj}^2 \tag{3.16}$$

where S_{kj} is the market share of firm k in industry j. The market share of an individual firm is calculated by using the firm's net sales divided by the total sales of the entire industry.[40] The calculation is performed every year, and the average value over the past three years is used as the HHI of an industry to prevent potential data errors in the analysis. A small value of HHI implies that the market is shared by many competing firms, while a large value suggests that the market share is concentrated in the hands of a few large firms.

We predict the news spillover effect to be weaker in a more concentrated product market for two reasons. First, negative news for a company can be potentially good news for its rivalry companies, and this is particularly true for industries dominated by a few firms. Second, bellwether firms' news becomes less influential in a more concentrated industry because industry peers receive more news coverage and the cost of information gathering is lower relative to those in less concentrated industries. If news from industry peers can also capture a similar common subset of information, news about bellwether firms becomes less important.

Results in Table 3.9 provide support for this prediction. In regression (1), we document a significant positive coefficient on $NewsTone_{IBW,t}$, and a significant negative coefficient on the interaction term $NewsTone_{IBW,t} * HHI_{k,t}$. It suggests the positive

[40] We classify industries using 48 Fama and French (1997) industry classifications, and all firms with no-missing sales data are included when calculating the Herfindahl index for a particular industry. Results are similar when industries are classified using three-digit SIC business segments.

relation between the bellwether firms' news tone and industry peers' returns is weaker in more concentrated industries. Similarly, the negative coefficient on the interaction term in regression (2) indicates that the positive association between bellwether firms' news tone and industry peers' analyst forecast revisions is less significant in more concentrated industries. Also, we find that bellwether firms' news coverage exhibits less influence on industry peers' trading activity in more concentrated industries (regression (3)). This finding supports the explanation that, in a more concentrated industry, investors are less dependent on the information from leading firms to make the trading decisions. Overall, results in Table 3.9 suggest that intra-industry news spillover is related to product market competition.

Table 3.9 News spillover and product market competition

This table presents the impact of industry bellwether firm news on industry peers' stock returns, trading activity and analyst forecasts conditioning on industry concentration. Industry concentration is measured by the Herfindahl-Hirschman Index (HHI), which is computed based on the market share of each firm in the industry. Regressions (1) and (2) report the impact of the news tone ($NewsTone$). Regressions (3) and (4) report the impact of the number of news ($NNews$). $R_{k,t}$ is firm *k*'s stock return in month *t*; $FR_{k,t}$ is the revision in consensus forecasts of 1-year ahead earnings per share, $VOSHOCK_{k,t}$ is the trading volume shock, and $FA_{k,t}$ is the analyst forecast accuracy, defined as the negative of the absolute value of the analyst forecast error. $NewsTone_{IBW}$ and $NNews_{IBW}$ is the news tone score and number of news for industry bellwether firms (i.e., same industry as firm *k*), respectively. $Controls$ represents control variables specified in Table 3.3, including firm *K*'s market capitalization ($Size_k$), book-to-market ratio (BM_k), average daily share turnover ($Turnover_k$), the fraction of shares outstanding held by institutional investors (IO_k), the lagged stock return ($R_{k,t-1}$), the cumulative return over month *t*-7 to *t*-2 ($R_{k,t-2,t-7}$), and value-weighted returns of all stocks in CRSP in month *t*-1 ($R_{m,t-1}$) and month *t* ($R_{m,t}$). The table reports the results from panel regressions with industry and month fixed effects. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $R_{k,t}$	(2) $FR_{k,t}$	(3) $VolShock_{k,t}$	(4) $FA_{k,t}$
$NewsTone_{IBW.t}$	0.0502***	0.0042**		
	(3.17)	(2.71)		
$NewsTone_{k,t}$	0.0353***	0.0019***		
	(15.02)	(5.81)		
$NNews_{IBW,t}$			0.0501**	0.0709**
			(2.58)	(2.23)
$NNews_{k,t}$			0.4022***	-0.0869***
			(17.76)	(-8.28)
$NewsTone_{IBW,t} * HHI_{k,t}$	-0.2809***	-0.0053*		
	(-3.28)	(-1.69)		
$NNews_{IBW,t} * HHI_{k,t}$			-0.0876*	-0.0815
			(-1.92)	(-1.42)
$HHI_{k,t}$	0.1210*	0.0002	0.2846	0.2131
	(2.01)	(0.34)	(1.62)	(1.19)
$Controls_{k,t-1}$	Yes	Yes	Yes	Yes
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	353,848	239,173	355,513	289,073
R-squared	0.1777	0.0456	0.1051	0.0379

3.6.2.2 News spillover in different market states

Existing literature implies that information production varies with market states. For example, Veldkamp (2006) suggests that information provision varies with asset values.[41] When asset values are low (high), information provision falls (rises). Veldkamp (2006)'s prediction on the positive relation between information provision and asset values can have two competing implications for information production in our study. First, increasing information provision makes more investors learn about the same subset of assets. Thus, it predicts a stronger news spillover effect in up markets when asset value rises. Second, increasing information provision makes information available for a broader class of assets, which reduces the extent to which investors use information about one asset to make inference about others. Thus, it predicts a weaker news spillover in up market. The two opposite predictions leave news spillover across market states an empirical question.

In addition, behavioral finance literature argues that investors are subject to sentiment, and market-wide investor sentiment has significant impact on market activities (e.g., Baker and Wurgler, 2007; De Long, Shleifer, Summers, and Waldmann, 1990). Using aggregate market performance as the proxy for market sentiment, several studies find that market reaction to news varies with aggregate market performance (Conrad, Cornell, and Landsman, 2002; Mian and Sankaraguruswamy, 2012). Motivated by the literature, we further explore whether the observed news spillover effect varies with market states.

Following Cooper, Gutierrez, and Hameed (2004), we define market state using the cumulative return of the value-weighted CRSP market index over the past 36 months. A month is labelled as an up-market month if the CRSP index return is positive, and as a down-market month if the CRSP index return is negative.[42] Over our sample period January 2003 – March 2016, there are 117 up months and 42 down months.

We augment Equations (3.8a) and (3.8b) with an interaction term between bellwether firm news and up market dummy:

$$R_{k,t}\ (or\ FR_{k,t}) = \alpha_0 + \beta_1 NewsTone_{IBW,t} + \beta_2 NewsTone_{IBW,t} * UP_t + \beta_3 UP_t + \beta_4 NewsTone_{k,t} + Other\ Control_{k,t-1} + \varepsilon_{k,t} \quad (3.17a)$$

[41] The cyclical variation in information flow is supported by numerous empirical studies (e.g., Brockman, Liebenberg, and Schutte ,2010; Ribeiro and Veronesi, 2002).

[42] We also consider a two-year and a one-year definition of the market's state and the results are qualitative similar.

$$VOSHOCK_{k,t}(or\ FA_{k,t}) = \alpha_0 + \beta_1 NNEWS_{IBW,t} + \beta_2 NNEWS_{IBW,t} * UP_t + \beta_3 UP_t + \beta_4 NNEWS_{k,t} + Other\ Control_{k,t-1} + \varepsilon_{k,t} \quad (3.17b)$$

where UP_t is a dummy variable taking a value of one if the market is defined as an up-market, and zero otherwise. The coefficient on interaction term, β_2 captures how market state affects the news spillover from bellwether firms to industry peers. We report the regression results in Table 3.10.

Coefficients on the interaction term are significant in regressions (2) and (3) when $FR_{k,t}$ and $VOSHOCK_{k,t}$ are used as dependent variables, respectively. Thus, our results provide evidence that information spillover varies across market states. The findings are also in line with the existing information production literature. The negative coefficient on the interaction term in regression (2) is consistent with the interpretation that, because information is abundant for a broader class of assets in up markets, it reduces the extent to which analysts use information about bellwether firms to make inference about other firms. The significant positive coefficient in regression (3) suggests that the impact of bellwether firms' news coverage on industry peers' trading volume is greater in up markets. The result is consistent with the investor sentiment explanation. Market reaction to news is stronger when market sentiment is high.

Table 3.10 News spillover across market states

This table presents the impact of industry bellwether firm news on industry peers' stock returnss, trading activity and analyst forecasts across different market states. Market state is measured by the cumulative return of the value-weighted CRSP market index over the past 36 months. We label a month as an up-market month if the CRSP index return is positive, and as a down-market month if the CRSP index return is negative. UP_t is a dummy variable taking a value of one if the market is defined as an up-market, and zero otherwise. Regressions (1) and (2) report the impact of the news tone ($NewsTone$). Regressions (3) and (4) report the impact of the number of news ($NNews$). $R_{k,t}$ is firm *k*'s stock return in month *t*; $FR_{k,t}$ is the revision in consensus forecasts of 1-year ahead earnings per share, $VOSHOCK_{k,t}$ is the trading volume shock, and $FA_{k,t}$ is the analyst forecast accuracy, defined as the negative of the absolute value of the analyst forecast error. $NewsTone_{IBW}$ and $NNews_{IBW}$ is the news tone score and number of news for industry bellwether firms (i.e., same industry as firm *k*), respectively. $Controls$ represents control variables specified in Table 3.3, including firm *k*'s market capitalization ($Size_k$), book-to-market ratio (BM_k), average daily share turnover ($Turnover_k$) , the fraction of shares outstanding held by institutional investors (IO_k), the lagged stock return ($R_{k,t-1}$), the cumulative return over month t-7 to t-2 ($R_{k,t-2,t-7}$), and value-weighted returns of all stocks in CRSP in month *t*-1 ($R_{m,t-1}$) and month *t* ($R_{m,t}$). The table reports the results from panel regressions with industry and month fixed effects. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $R_{k,t}$	(2) $FR_{k,t}$	(3) $VOSHOCK_{k,t}$	(4) $FE_{k,t}$
$NewsTone_{IBW.t}$	0.0290***	0.0051**		
	(2.97)	(2.50)		
$NewsTone_{IBW,t} * UP_t$	-0.0033	-0.0052**		
	(-0.26)	(-2.44)		
$NewsTone_{k,t}$	0.0354***	0.0017***		
	(15.06)	(3.87)		
$NNews_{IBW,t}$			-0.0122	0.0505**
			(-1.29)	(2.05)
$NNews_{IBW,t} * UP_t$			0.0666***	0.0161
			(4.86)	(0.93)
$NNews_{k,t}$			0.4038***	-0.0749***
			(12.09)	(-6.95)
UP_t	-0.0027	-0.0083***	0.4404***	-0.0998
	(-0.24)	(-4.23)	(5.04)	(-1.50)
$Controls_{k,t-1}$	Yes	Yes	Yes	Yes
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	353,848	239,204	355,568	289,076
R-squared	0.1776	0.0454	0.1058	0.0380

3.6.2.3 News spillover and the sign of news tone

Prior studies suggest that due to behavioural biases, the market responses to news is asymmetric, and market reacts more strongly to bad news than to good news (Barberis, Shleifer, and Vishny, 1998; Skinner and Sloan, 2002). Hou (2007) shows that stock prices of small firms react more strongly to negative news of the big firms in the same industry. Building upon the literature, we expect negative news of bellwether firms to have stronger influence on their industry peers than the positive news. To test this conjecture, we set a negative news dummy, $Negative_{IBM,t}$, as one if the average news tone of the bellwether firms in that industry is negative in month t, and zero otherwise.[43] In Table 3.11, we employ the same regressions as those in Table 3.3 but further include an interaction between bellwether firm news and negative news dummy.

Our results are consistent with the previous findings that market reacts more strongly to negative news. Regression (1) shows a positive coefficient on the interaction term, $NewsTone_{IBW,t} * Negative_{IBM,t}$, suggesting industry peers' price reaction is stronger when the bellwether firms' news is negative. The coefficient on the interaction term is also significantly positive in regression (2), when testing the impact of bellwether firm news on the analyst forecast revision on industry peers. Noticeably, the significant relation between $NewsTone_{IBW,t}$ and $FR_{k,t}$ observed in Table 3.3 disappears after including the interaction term. This suggests the influence of bellwether firm news tone on industry peers' analyst forecast is driven by the negative news. In addition, we find that the effect of bellwether firms' news coverage on industry peers' trading volume is stronger when the news is negative. Coefficient on the interaction term, $NNews_{IBW,t} * Negative_{IBM,t}$, is positive and statistically significant in regression (3).

[43] There are 688 negative observations out of 5857 industry-months in our sample.

Table 3.11 News spillover and negative news tone

This table presents the impact of industry bellwether firm news on industry peers' stock prices, trading activity and analyst forecasts considering the sign of the news tone. $Negative_{IBM,t}$ is a dummy variable taking a value of one if the average news tone on bellwether firms is negative in month t, and zero otherwise. $R_{k,t}$ is firm *k*'s stock return in month t; $FR_{k,t}$ is the revision in consensus forecasts of 1-year ahead earnings per share, $VOSHOCK_{k,t}$ is the trading volume shock, and $FA_{k,t}$ is the analyst forecast accuracy, defined as the negative of the absolute value of the analyst forecast error. $NewsTone_{IBW}$ and $NNews_{IBW}$ is the news tone score and number of news for industry bellwether firms (i.e., same industry as firm *k*), respectively. $Controls$ represents control variables specified in Table 3.3, including firm *k*'s market capitalization ($Size_k$), book-to-market ratio (BM_k), average daily share turnover ($Turnover_k$), the fraction of shares outstanding held by institutional investors (IO_k), the lagged stock return ($R_{k,t-1}$), the cumulative return over month *t*-7 to *t*-2 ($R_{k,t-2,t-7}$), and value-weighted returns of all stocks in CRSP in month *t*-1 ($R_{m,t-1}$) and month t ($R_{m,t}$). The table reports the results from panel regressions with industry and month fixed effects. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $R_{k,t}$	(2) $FR_{k,t}$	(3) $VOSHOCK_{k,t}$	(4) $FE_{k,t}$
$NewsTone_{IBW.t}$	0.0140**	0.0004		
	(2.66)	(0.69)		
$NewsTone_{IBW,t} * Negative_{IBM,t}$	0.1097*	0.0037**		
	(1.73)	(3.05)		
$NewsTone_{k,t}$	0.0373***	0.0014***		
	(13.57)	(6.37)		
$NNews_{IBW,t}$			0.0236*	0.0511**
			(1.84)	(2.08)
$NNews_{IBW,t} * Negative_{IBM,t}$			0.0664**	0.0065
			(3.17)	(0.29)
$NNews_{k,t}$			0.4036***	-0.0872***
			(12.06)	(-8.31)
$Negative_{IBM,t}$			-0.2091**	0.0638
			(-3.19)	(0.77)
$Controls_{k,t-1}$	Yes	Yes	Yes	Yes
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	353,848	239,204	355,568	289,076
R-squared	0.1320	0.0534	0.1058	0.0380

3.6.2.4 News spillover in different news coverage groups

Previous studies suggest that the intra-industry information spillover effect may vary across firms depending on their information environment. For example, Hameed et al. (2015) find that the effect of the analyst earning forecast revisions for bellwether firms decreases monotonically from the least to most analyst covered firms. Hou (2007) and Menzly and Ozbas (2010) show that information tends to diffuse from firms with less information uncertainty to firms with more information uncertainty. Motivated by those studies, we examine the impact of bellwether firm news on industry peers with different quality of information environment. We use news coverage as the measure of a firm's information opaqueness.[44] For each month and each industry, we group stocks into terciles based on the number of news coverage. We examine news spillover effect in separate news coverage groups and present the results in Table 3.12.

In line with the prior studies, we find that the news spillover is stronger among industry peers with more opaque information environment. Although bellwether firms' news tone is associated with stock price movements of industry peers from all three news coverage groups (Panel A), it only exhibits significant impact on the analyst forecast revisions for peer firms with low and medium news coverage (Panel B). Panel C shows that bellwether firms' news coverage only affects the trading volume for stocks from the low media coverage group. Likewise, Panel D indicates that bellwether firms' news coverage improves the analyst forecast accuracy for industry peers that are less followed by the news media. This finding is consistent with the view that investors use information about informative stocks to infer the value of more opaque stocks.

[44] We also use the firm's institutional ownership and analyst coverage as the proxy for information opaqueness. The results are qualitatively similar.

Table 3.12 News spillover conditioned on industry peers' news coverage

This table shows the information spillover from bellwether firms to industry peers from different news coverage groups. For each month and each industry, we sort stocks (exclude bellwether stocks) into three groups based on the news coverage. Low (Medium and High) coverage group contains stocks from the bottom (medium and high) news coverage tercile. Panels A to D report the impact of bellwether firm news on stocks returns, analyst forecast revisions, trading volume and analyst forecast accuracy for industry peers from different news coverage groups, respectively. $R_{k,t}$ is firm k's stock return in month t; $FR_{k,t}$ is the revision in consensus forecasts of 1-year ahead earnings per share, $VOSHOCK_{k,t}$ is the trading volume shock, and $FA_{k,t}$ is the analyst forecast accuracy, defined as the negative of the absolute value of the analyst forecast error. $NewsTone_{IBW}$ and $NNews_{IBW}$ is the news tone score and number of news for industry bellwether firms (i.e., same industry as firm k), respectively. $Controls$ represents control variables specified in Table 3.3, including firm k's market capitalization ($Size_k$), book-to-market ratio (BM_k), average daily share turnover ($Turnover_k$) , the fraction of shares outstanding held by institutional investors (IO_k), the lagged stock return ($R_{k,t-1}$), the cumulative return over month t-7 to t-2 ($R_{k,t-2,t-7}$), and value-weighted returns of all stocks in CRSP in month t-1 ($R_{m,t-1}$) and month t ($R_{m,t}$). The table reports the results from panel regressions with industry and month fixed effects. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	Low coverage	*Medium coverage*	*High coverage*
Panel A: Bellwether firm news tone and industry peers' stock return			
	$R_{k,t}$	$R_{k,t}$	$R_{k,t}$
$NewsTone_{IBW,t}$	0.0176*	0.0283***	0.0287**
	(1.84)	(3.14)	(2.61)
$NewsTone_{k,t}$	0.0163***	0.0441***	0.0911***
	(12.47)	(24.81)	(15.09)
$Controls_{k,t-1}$	Yes	Yes	Yes
Observations	117,637	117,787	118,424
R-squared	0.1742	0.1836	0.1942
Panel B: Bellwether firm news tone and industry peers' forecast revision			
	$FR_{k,t}$	$FR_{k,t}$	$FR_{k,t}$
$NewsTone_{IBW,t}$	0.0023**	0.0034**	0.0025
	(2.68)	(2.54)	(1.53)
$NewsTone_{k,t}$	0.0007***	0.0027***	0.0035***
	(2.95)	(5.70)	(6.27)
$Controls_{k,t-1}$	Yes	Yes	Yes
Observations	68,040	80,035	91,129
R-squared	0.0406	0.0478	0.0542
Panel C: Bellwether firm news coverage and industry peers' trading volume			
	$VOSHOCK_{k,t}$	$VOSHOCK_{k,t}$	$VOSHOCK_{k,t}$
$NNews_{IBW,t}$	0.0290*	0.0246	0.0294
	(1.70)	(1.59)	(1.27)
$NNews_{k,t}$	0.1673***	0.3692***	0.6543***
	(11.80)	(10.35)	(11.71)
$Controls_{k,t-1}$	Yes	Yes	Yes
Observations	118,050	118,183	119,335
R-squared	0.0836	0.0877	0.1413
Panel D: Bellwether firm news coverage and industry peers' analyst forecast accuracy			
	$FE_{k,t}$	$FE_{k,t}$	$FE_{k,t}$
$NNews_{IBW,t}$	0.1241***	0.0536*	0.0148
	(2.97)	(1.79)	(0.58)
$NNews_{k,t}$	-0.0822**	-0.0723**	-0.1364***
	(-2.36)	(-2.01)	(-5.25)
$Controls_{k,t-1}$	Yes	Yes	Yes
Observations	80,516	97,174	112,684
R-squared	0.0390	0.0418	0.0410

3.6.3 News spillover among bellwether firms

Having established the evidence on the news spillover from bellwether firms to their industry peers, we further investigate whether news spillover exists among the bellwether firms themselves. We conduct the similar analysis as in Table 3.3 to exam how an individual bellwether firm's stock prices, trading activity and analyst forecasts are affected by the news on the other bellwether firms in the same industry. Results are presented in Table 3.13. We find that a bellwether firm's stock prices and analyst forecast revisions are positively affected by the news tone of its industry bellwether counterparts, coefficients on $NewsTone_{Peer\_IBW}$ are positive and statistically significant at the 1% level in both regressions (1) and (2). Regression (3) suggests that bellwether firms' news coverage also exhibits significant impact on the trading activity of the other bellwether firms with the same industry. In regression (4), we find that a bellwether firm's analyst forecast accuracy is not significantly affected by its peer bellwether firms' news coverage. Unlike more opaque stocks, analysts do not reply on information about other leading firms to infer the changes in fundamental value for an industry bellwether firm.

3.6.4 News partial correlation and return comovement across market states

Extant studies suggest that time-varying information production drives the comovement patterns over time. Veldkamp's (2006) theoretical model predicts that information production increases during economic expansion and decreases during contraction, and information production can help explain fluctuations in stock return comovement. Consistent with this prediction, Hochstotter, Meyer, Riordan, and Storkenmaier (2014) show that news production helps explain time series variation in country-level stock market comovement. Brockman et al. (2010) document countercyclical pattern of return comovement and establish a causal relation from information production over business cycles to stock return comovement. Motivated by the evidence on the link between macro economy, information production, and return comovement, we investigate whether aggregate market performance plays a role affecting the relation between news partial correlation and return comovement.

Table 3.14 reports the results from regressions of return comovement on news partial correlation conditional on the market states. Consistent with the findings in Table 3.6, we document a positive relation between news partial correlation and return comovement. Coefficients on the interaction term $LPCORR\_NNEWS_{k,t} * UP_t$ in regressions (3) and (4) are significantly positive, suggesting that the association between news partial correlation and return comovement is stronger in up markets. Moreover, in line with Brockman et al. (2010) who document countercyclicality in comovement, we show that return comovement is negatively associated with aggregate market performance. Coefficients on the up-market dummy (UP_t) are negative and statistically significant at the 1% level in all four regressions.

Table 3.13 News spillover in bellwether firms

This table presents the news spillover among industry bellwether firms. For each year and each industry, we first sort stocks into terciles based on analyst coverage. Within the top analyst coverage tercile, we further rank stocks into three groups based on their $LPCORR\_ROA$. Stocks with high $LPCORR\_ROA$ among those high analyst coverage stocks are defined as bellwether firms. Regressions (1) and (2) investigate how a bellwether firm's returns and analyst forecasts revisions are affected by news tone of other bellwether firms within the same industry ($NewsTone_{Peer\_IBW,t}$). Regressions (3) and (4) examine how the bellwether firm's trading volume and analyst forecast accuracy is influenced by the news coverage of its peer industry bellwether firms ($NNews_{Peer\_IBW,t}$). For industry bellwether firm *i*, $R_{i,t}$ is firm *i*'s stock return in month *t*. $FR_{i,t}$ is the revision in consensus forecasts of 1-year ahead earnings per share, $VOSHOCK_{i,t}$ is the trading volume shock, and $FA_{i,t}$ is the analyst forecast accuracy, defined as the negative of the absolute value of the analyst forecast error. $NewsTone_{Peer\_IBW}$ and $NNews_{Peer\_IBW}$ is the average news tone score and number of news across firm *i*'s industry bellwether firm peers (excluding firm *i*), respectively. $Controls$ represents control variables specified in Table 3.3, including firm *k*'s market capitalization ($Size_k$), book-to-market ratio (BM_k), average daily share turnover ($Turnover_k$), the fraction of shares outstanding held by institutional investors (IO_k), the lagged stock return ($R_{k,t-1}$), the cumulative return over month t-7 to t-2 ($R_{k,t-2,t-7}$), and value-weighted returns of all stocks in CRSP in month t-1 ($R_{m,t-1}$) and month t ($R_{m,t}$). The table reports the regression results across firm-months with industry and month fixed effects. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $R_{i,t}$	(2) $FR_{i,t}$	(3) $VolShock_{i,t}$	(4) $FA_{i,t}$
$NewsTone_{Peer\_IBW,t}$	0.0265***	0.0017***		
	(3.19)	(3.06)		
$NewsTone_{i,t}$	0.0419***	0.0013***		
	(13.33)	(7.28)		
$NNews_{Peer\_IBM,t}$			0.0901***	-0.0029
			(3.34)	(-0.76)
$NNews_{i,t}$			0.3737***	-0.0158***
			(13.89)	(-3.34)
$Controls_{k,t-1}$	Yes	Yes	Yes	Yes
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	37,250	31,415	37,225	36,871
R-squared	0.2423	0.1453	0.2347	0.0477

Table 3.14 News partial correlation and return comovement in different market states

This table shows the relation between return comovement and news partial correlation in different market states. Market state is defined using the cumulative return of the value-weighted CRSP market index over the past 36 months. We label a month as an up-market month if the CRSP index return is positive, and as a down-market month if the CRSP index return is negative. UP_t is a dummy variable taking a value of one if the market is defined as an up-market, and zero otherwise. Dependent variable $RetComove_{k,t}$ represents the return comovement for stock *k* in month *t*, measured using the R^2 from regression of daily stock returns of stock *k* on equal-weighted market returns and taking the logarithm transformation. $LPCORR\_TONE$, $LPCORR\_NNEWS$ are news partial correlation measures, based on news tone score and the number of news, respectively. $Controls$ represent control variables specified in Table 3.6, including the firm's partial correlation in fundamentals ($LPCORR\_ROA$), market capitalization ($Size$), book-to-market ratio (BM), share price($Price$), monthly stock return ($R_{k.t-1}$) and the cumulative return over the past 6 months ($R_{k,t-2,t-7}$), return standard deviation ($RetStd$), the Amihud (2002) illiquidity ratio ($Liquidity$), the number of estimates making the one-year-ahead earnings forecasts ($Analyst$), and the fraction of shares outstanding held by institutional investors(IO). The table reports the panel regression results across all the firms. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $RetComove_{k,t}$	(2) $RetComove_{k,t}$	(3) $RetComove_{k,t}$	(4) $RetComove_{k,t}$
$LPCORR\_TONE_{k,t}$	0.0841***	0.0795***		
	(2.87)	(2.72)		
$LPCORR\_TONE_{k,t} * UP_t$	0.0377	0.0382		
	(1.18)	(1.20)		
$TONEComove_{k,t}$		0.0172***		
		(6.94)		
$LPCORR\_NNEWS_{k,t}$			0.0926***	0.0791***
			(3.38)	(2.89)
$LPCORR\_NNEWS_{k,t} * UP_t$			0.0623**	0.0604*
			(2.01)	(1.94)
$NNEWSComove_{k,t}$				0.0230***
				(7.78)
UP_t	-0.4505***	-0.4505***	-0.4037***	-0.4076***
	(-6.02)	(-6.03)	(-5.78)	(-5.84)
$Controls_{k,t-1}$	Yes	Yes	Yes	Yes
Observations	108,745	108,742	108,133	108,131
R-squared	0.1291	0.1298	0.1263	0.1275

3.7 Robustness tests

3.7.1 An alternative frequency for identifying bellwether firms

In the main analysis, we identify industry bellwether firms at the end of each year based on the stock's analyst coverage in December and the average *LPCORR_ROA* during the year. In this section, we identify bellwether firms on a quarterly basis. Following the same procedure, in each quarter, we first sort stocks from each industry into terciles using the average monthly analyst coverage over the quarter. Within the top analyst coverage tercile, we further rank stocks into three equal groups based on their quarterly *LPCORR_ROA*. Stocks from the top *LPCORR_ROA* group among those high analyst coverage stocks are defined as industry bellwether firms. We reproduce Table 3.3 using the quarterly identified bellwether firms and report the results in Table 3.15.

Results are comparable to those in Table 3.3. Panel A shows that coefficients on bellwether firms' news remain positive and statistically significant at the 1% level in all specifications. Panel B shows that the influence of bellwether firms' news remains after controlling for industry peers' own news. Thus, our findings on the news spillover from bellwether firms to industry peers are robust using alternative frequency for identifying bellwether firms.

Table 3.15 News of bellwether firms and its impact on industry peers - Identifying bellwether firms on a quarterly basis

This table presents the impact of industry bellwether firm news on industry peers' stock prices, trading activity and analyst forecasts. For each quarter and each industry, we first sort stocks into terciles based on analyst coverage. Within the top analyst coverage tercile, we further rank stocks into three groups based on their $LPCORR\_ROA$. Stocks with high $LPCORR\_ROA$ among those high analyst coverage stocks are defined as bellwether firms. Regressions (1) and (2) report the impact of bellwether firms' news tone ($NewsTone_{IBW,t}$). Regressions (3) and (4) report the impact of the number of news ($NNews_{IBW,t}$). For all firm k, excluding the industry bellwether firms, $R_{k,t}$ is firm k's stock return in month t. $FR_{k,t}$ is the revision in consensus forecasts of 1-year ahead earnings per share, $VOSHOCK_{k,t}$ is the trading volume shock, and $FA_{k.t}$ is the analyst forecast accuracy, defined as the negative of the absolute value of the analyst forecast error. $NewsTone_{IBW}$ and $NNews_{IBW}$ is the news tone score and number of news for industry bellwether firms (i.e., same industry as firm k), respectively. All control variables are defined in Equations (3.8a) and (3.8b). In Panel B, we add the news tone ($NewsTone_{k,t}$) and number of news ($NNews_{k,t}$)for firm k. The table reports the regression results across firm-months with industry and month fixed effects. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $R_{k,t}$	(2) $FR_{k,t}$	(3) $VolShock_{k,t}$	(4) $FA_{k,t}$
Panel A: Impact of bellwether firm news on other firms				
$NewsTone_{IBW,t}$	0.0295***	0.0023***		
	(3.14)	(3.55)		
$NNews_{IBW,t}$			0.0630***	0.0348**
			(4.22)	(2.07)
$R_{k,t-1}$	-0.0199***	0.0144***	0.5038***	0.5564***
	(-4.48)	(10.65)	(4.04)	(8.51)
$R_{k,t-2,t-7}$	0.0017	0.0038***	0.2929***	0.3363***
	(1.65)	(6.17)	(8.92)	(8.50)
$Size_{k,t-1}$	-0.0013***	0.0002***	-0.0212***	0.1000***
	(-5.89)	(3.68)	(-6.52)	(8.75)
$BM_{k,t-1}$	0.0017***	-0.0002***	0.0315***	-0.0464*
	(3.87)	(-3.35)	(4.01)	(-1.65)
$Turnover_{k,t-1}$	-0.2255***	-0.0852***	11.8041***	-12.2052***
	(-3.41)	(-5.51)	(10.58)	(-4.00)
$IO_{k,t-1}$	0.0032***	0.0001**	-0.0074**	0.1259***
	(8.38)	(2.39)	(-2.46)	(8.54)
$R_{m,t-1}$	-0.2702	0.0842**	-14.2535***	0.6368
	(-0.99)	(2.51)	(-5.43)	(0.99)
$R_{m,t}$	1.2989***	-0.1011***	11.0134***	-1.2090***
	(6.53)	(-4.55)	(9.02)	(-2.81)
Constant	0.0010	0.0068***	-0.4810***	-0.4709***
	(0.10)	(5.06)	(-7.59)	(-5.21)
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	454,112	281,290	455,320	339,185
R-squared	0.1649	0.0510	0.0605	0.0372
Panel B: Impact of bellwether firm news on other firms controlling for the firm's news				
$NewsTone_{IBW,t}$	0.0269***	0.0021***		
	(2.79)	(3.25)		
$NewsTone_{k,t}$	0.0352***	0.0016***		
	(14.74)	(10.08)		
$NNews_{IBW,t}$			0.0521***	0.0377**
			(3.94)	(2.08)
$NNews_{k,t}$			0.3294***	-0.0708***
			(17.86)	(-8.01)
$Controls_{k.t-1}$	Yes	Yes	Yes	Yes
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	354,475	239,754	354,766	288,762
R-squared	0.1781	0.0531	0.1026	0.0376

3.7.2 An alternative definition of bellwether firms

By design, bellwether stocks in our sample are heavily analysed stocks whose fundamentals best predict those of their industry peers. Thus, the observed information spillover effect might be purely driven by the correlations in fundamentals.[45] In this section, we redefine highly followed firms whose news contributes most to the news of their industry peers as bellwether firms. Each year, we first sort stocks from each industry into terciles based on their analyst coverage. Within the top analyst coverage tercile, we further rank stocks into three equal groups based on their partial correlation in news with industry peers ($LPCORR\_TONE$ or $LPCORR\_NNEWS$). Stocks from the top news partial correlation group among those high analyst coverage stocks are defined as industry bellwether firms.[46]

We reproduce Table 3.3 using the newly identified bellwether firms, and report the results in Table 3.16. Panel A presents the news spillover from bellwether firms identified based on $LPCORR\_TONE$. Panel B shows the bellwether firms identified based on $LPCORR\_NNEWS$. Results are similar as those reported in Table 3.3. The positive impact of bellwether firms' news on industry peers remains in all regressions. Therefore, our findings are robust using alternative bellwether firm identification criteria. More importantly, the robust news spillover implies that the observed intra-industry information transmission may not be fully explained by fundamentals.

[45] Although this is highly unlikely, given the week positive correlation between news partial correlation and partial correlation in fundamentals reported in Table 1.
[46] There is a moderate overlap in the selected bellwether firms using the two different ways of identifying industry bellwether firms. Among the 3882 (3876) bellwether firm-years based on $LPCORR\_TONE$ ($LPCORR\_NNEWS$), 1245 (1239) firm-years overlap with the observations based on $LPCORR\_ROA$.

Table 3.16 News of bellwether firms and its impact on industry peers - Identifying bellwether firms based on news partial correlation

This table presents the impact of industry bellwether firm news on industry peers' stock prices, trading activity and analyst forecasts, where bellwether firms are identified based on news coverage and partial correlation in news. For each year and each industry, we first sort stocks into terciles based on analyst coverage. Within the top analyst coverage tercile, we further rank stocks into three groups based on their $LPCORR\_TONE$ (Panel A) or $LPCORR\_NNEWS$ (Panel B). Stocks with high $LPCORR\_TONE$ (or $LPCORR\_NNEWS$) among those high analyst coverage stocks are defined as bellwether firms. Regressions (1) and (2) report the impact of bellwether firms' news tone ($NewsTone_{IBW,t}$). Regressions (3) and (4) report the impact of the number of news ($NNews_{IBW,t}$). For all firm k, excluding the industry bellwether firms, $R_{k,t}$ is firm k's stock return in month t. $FR_{k,t}$ is the revision in consensus forecasts of 1-year ahead earnings per share, $VOSHOCK_{k.t}$ is the trading volume shock, and $FA_{k,t}$ is the analyst forecast accuracy, defined as the negative of the absolute value of the analyst forecast error. $NewsTone_{IBW}$ and $NNews_{IBW}$ is the news tone score and number of news for industry bellwether firms (i.e., same industry as firm k), respectively. All control variables are defined in Equations (3.8a) and (3.8b). In Panel B, we add the news tone ($NewsTone_{k,t}$) and number of news ($NNews_{k,t}$)for firm k. The table reports the regression results across firm-months with industry and month fixed effects. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $R_{k,t}$	(2) $FR_{k,t}$	(3) $VolShock_{k,t}$	(4) $FA_{k,t}$
Panel A: Bellwether firms identified based on $LPCORR\_TONE$				
$NewsTone_{IBW,t}$	0.0146*	0.0009**		
	(1.71)	(2.12)		
$NewsTone_{k,t}$	0.0355***	0.0017***		
	(23.83)	(9.61)		
$NNews_{IBW,t}$			0.0239*	0.0311**
			(1.81)	(2.12)
$NNews_{k,t}$			0.3994***	-0.0932***
			(26.23)	(-7.08)
$R_{k,t-1}$	-0.0154***	0.0145***	0.5851***	0.5726***
	(-4.93)	(9.87)	(6.37)	(4.92)
$R_{k,t-2,t-7}$	0.0010	0.0039***	0.3919***	0.3142***
	(0.91)	(6.00)	(14.68)	(6.15)
$Size_{k,t-1}$	-0.0016***	0.0002***	-0.1783***	0.1252***
	(-7.97)	(3.21)	(-25.03)	(6.48)
$BM_{k,t-1}$	0.0026***	-0.0002***	-0.0027	-0.0559**
	(6.55)	(-3.70)	(-0.39)	(-2.40)
$Turnover_{k,t-1}$	-0.1936***	-0.0901***	4.6966***	-10.1488***
	(-3.53)	(-4.89)	(6.14)	(-3.24)
$IO_{k,t-1}$	0.0029***	0.0001*	0.0166***	0.1272***
	(8.52)	(1.83)	(5.06)	(6.33)
$R_{m,t-1}$	-0.3874*	0.0890***	-13.1239***	-0.8655
	(-1.78)	(2.76)	(-8.33)	(-1.19)
$R_{m,t}$	1.4094***	-0.0944***	10.1606***	-0.6569
	(10.66)	(-4.42)	(11.30)	(-1.07)
Constant	-0.0031	0.0067***	0.0063	-0.5305***
	(-0.38)	(5.63)	(0.08)	(-4.61)
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	355,796	240,919	356,132	289,388
R-squared	0.1785	0.0559	0.1048	0.0382

Table 3.16 continued

Panel B: Bellwether firms identified based on $LPCORR\_NNEWS$				
$NewsTone_{IBW,t}$	0.0147*	0.0019**		
	(1.95)	(2.22)		
$NewsTone_{k,t}$	0.0355***	0.0021***		
	(23.79)	(7.42)		
$NNews_{IBW,t}$			0.0292**	0.0330**
			(2.17)	(2.29)
$NNews_{k,t}$			0.3997***	-0.0933***
			(25.63)	(-7.10)
$Controls_{k,t-1}$	Yes	Yes	Yes	Yes
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	355,413	240,759	355,745	289,085
R-squared	0.1784	0.0451	0.1045	0.0382

3.7.3 An alternative industry classification

We conduct the main intra-industry analysis based on 48 Fama and French (1997) industry classifications. As a robustness check, we adopt an alternative industry classification method – text-based network industry classification (TNIC) developed by Hoberg and Phillips (2010,2016). TNIC is based on the premise that product similarity is core to classifying industries and is constructed using the firm pairwise similarity scores from text analysis of firm 10K product descriptions. In this classification system, each firm has its own set of time-varying distinct competitors.[47]

We follow a similar procedure and identify the bellwether stocks based on analyst coverage and $LPCORR\_ROA$. Due to the nature of TNIC, bellwether stocks are identified within each individual firm's product market. Each year, we sort firm k's competitors into terciles based on their analyst coverage. Within the top analyst coverage tercile, we further rank competitors into three equal groups based on their partial correlation in fundamentals ($LPCORR\_ROA$). Firms with high $LPCORR\_ROA$ among those high analyst coverage firms are identified as the bellwether firms in firm k's product market. We examine how firm k is affected by the news on the leading firms in the same product market and present the results in Table 3.17.

Consistent with the findings from the main test, in Panel A, regressions (1), (2) and (3) show that bellwether firms' news has significant impact on stock prices, analyst forecast revisions and trading activity of the other firms in the same industry. TNIC not only provides the related industry firms for each stock, but also calculates firm-by-firm

[47] The authors show that TNIC improves upon SIC and NAICS codes in explaining differences in key characteristics across industries.

pairwise similarity scores, which is a real number in the interval [0,1] describing how similar the products of the two firms are. A high score indicates the two firms are near rivals. We utilize the similarity score to investigate how a firm responses to the news on their nearest rivals. Each year, we sort the firm's competitors by similarity scores and the one with the highest score is nominated as the firm's nearest rival.[48] In panel B, we augment regressions with the news on the firm's nearest rival. Results show that a stock is influenced by the news of its nearest competitor in a similar pattern as by the news of industry bellwether firms. However, the magnitude of the impact is not as strong as the bellwether firms. For example, in regression (1), coefficient on $NewsTone_{COM,t}$ is 0.0077 which is much smaller compared to the coefficient of 0.0265 on $NewsTone_{IBW\_TNIC}$. We argue that negative firm-specific news of a close competitor sometimes can be good news for the firm. The confounding effect weakens the positive association between the firm's stock prices and its close competitors news tone. Overall, results from Table 3.17 suggest that our findings on the news spillover remain robust using alternative industry classifications.

[48] We exclude the bellwether firms when selecting the nearest rivals.

Table 3.17 News of bellwether firms and its impact on industry peers - Using text-based network industry classifications

This table presents the impact of industry bellwether firm news on industry peers' stock prices, trading activity and analyst forecasts, where industry groups are identified using the text-based network industry classifications (TNIC) developed by Hoberg and Philips (2010). For each year and each stock, we first sort the pairwise industry firms in into terciles based on analyst coverage. Within the top analyst coverage tercile, we further rank stocks into three groups based on their $LPCORR\_ROA$. Stocks with high $LPCORR\_ROA$ among those high analyst coverage stocks are defined as bellwether firms. Regressions (1) and (2) report the impact of bellwether firms' news tone ($NewsTone_{IBW\_TNIC,t}$). Regressions (3) and (4) report the impact of the number of news ($NNews_{IBW\_TNIC,t}$). For all firm k, excluding the industry bellwether firms, $R_{k,t}$ is firm k's stock return in month t. $FR_{k,t}$ is the revision in consensus forecasts of 1-year ahead earnings per share, $VOSHOCK_{k,t}$ is the trading volume shock, and $FA_{k,t}$ is the analyst forecast accuracy, defined as the negative of the absolute value of the analyst forecast error. $NewsTone_{IBW\_TNIC}$ and $NNews_{IBW\_TNIC}$ is the news tone score and number of news for firm k 's industry bellwether firms, respectively. All control variables are defined in Equations (3.8a) and (3.8b). In Panel B, we add the news tone ($NewsTone_{COM,t}$) and number of news ($NNews_{COM,t}$) for firm *k's* nearest industry rival, who has the highest pairwise product similarity scores with firm k . The table reports the regression results across firm-months with firm and month fixed effects. T-statistics based on firm-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $R_{k,t}$	(2) $FR_{k,t}$	(3) $VolShock_{k,t}$	(4) $FA_{k,t}$
Panel A: News spillover from industry bellwether firms				
$NewsTone_{IBW\_TNIC,t}$	0.0247***	0.0018***		
	(9.80)	(3.59)		
$NewsTone_{k,t}$	0.0454***	0.0023***		
	(40.71)	(8.89)		
$NNews_{IBW\_TNIC,t}$			0.0289***	0.0181
			(4.40)	(1.24)
$NNews_{k,t}$			0.0093***	-0.0351***
			(7.00)	(-4.00)
$R_{k,t-1}$	-0.0137***	0.0189***	0.4835***	0.1961***
	(-4.69)	(18.46)	(14.76)	(3.21)
$R_{k,t-2,t-7}$	0.0016**	0.0038***	0.2232***	0.0272
	(2.39)	(10.13)	(6.56)	(1.05)
$Size_{k,t-1}$	-0.0320***	0.0016***	-0.1088***	0.4143***
	(-37.62)	(4.55)	(-8.40)	(7.34)
$BM_{k,t-1}$	-0.0004	-0.0000	0.0720***	-0.0135
	(-0.56)	(-0.07)	(6.62)	(-0.43)
$Turnover_{k,t-1}$	-0.3412***	-0.1379***	13.8474***	1.8136
	(-6.18)	(-8.75)	(21.26)	(0.54)
$IO_{k,t-1}$	0.0042***	0.0001	0.0003	-0.1483***
	(9.12)	(0.38)	(0.06)	(-3.44)
$R_{m,t-1}$	-1.0168***	0.2231***	-17.7848***	7.6903***
	(-9.50)	(6.95)	(-14.94)	(4.68)
$R_{m,t}$	1.9518***	-0.1578***	12.8630***	-7.4581***
	(29.07)	(-5.31)	(16.66)	(-5.91)
Constant	0.1575***	0.0050**	-0.0656	-2.6815***
	(26.37)	(2.03)	(-0.78)	(-7.88)
Firm fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	319,986	221,133	320,741	265,393
R-squared	0.1950	0.0425	0.0906	0.0273

Table 3.17 continued

Panel B: News spillover from industry bellwether firms and the competitor				
$NewsTone_{IBW\_TNIC,t}$	0.0265***	0.0017***		
	(9.72)	(3.33)		
$NewsTone_{COM,t}$	0.0077***	0.0013***		
	(7.24)	(4.93)		
$NewsTone_{k,t}$	0.0480***	0.0022***		
	(39.70)	(8.27)		
$NNews_{IBW\_TNIC,t}$			0.0268***	0.0191
			(3.87)	(1.24)
$NNews_{COM,t}$			0.0198***	-0.0001
			(4.86)	(-0.01)
$NNews_{k,t}$			0.0090***	-0.0347***
			(6.79)	(-3.90)
$Controls_{k,t-1}$	Yes	Yes	Yes	Yes
Firm fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	281,745	196,622	282,426	235,765
R-squared	0.1979	0.0410	0.0926	0.0268

3.7.4 News correlation and return comovement with industry and market

We also measure return comovement by regressing individual stock returns on both industry and market returns. For each firm-month, we estimate the following regression using daily returns over a three-month window to obtain R^2 :

$$r_{k,t} = \alpha_k + \beta_{1k} r_{I,t} + \beta_{2k} r_{M,t} + \varepsilon_{k,t} \qquad (3.18)$$

where $r_{k,t}$ is the return of stock k on day t. $r_{I,t}$ is the industry return, computed as the equally weighted return for the industry, defined using 48 Fama and French (1997) industry classifications, on day t (excluding stock k). $r_{M,t}$ is the equally weighted market return on day t (excluding stock k). Return comovement is estimated by taking the log transformation of the regression R^2 as in Equation (3.13).

We rerun Equation (3.14) using this alternative return comovement measure and present the results in Table 3.18. Regression coefficients are much like those reported in Table 3.6. News partial correlation is positively associated with contemporaneous (Panel A) and future return comovement (Panel B). Regression R^2s are noticeably larger than those in Table 3.6. For example, R^2 is equal to 41.28% for regression (1) in Panel A, compared to 31.70% in Table 3.6. This is because partial correlation in news captures the intra-industry information flows, hence, it helps explain the return comovement at the industry level.

Table 3.18 Partial correlation in news and stock return comovement - An alternative measure for return comovement

This table presents results from panel regression of stock return comovement on partial correlation in news. Dependent variable $RetComove_{k,t}$ represents the return comovement for stock k in month t, measured using the R^2 from regression of daily stock returns of stock k on equal-weighted industry and market returns and taking the logarithmic transformation. $LPCORR_{TONE}$ and $LPCORR\_NNEWS$ are news partial correlation measures, based on news tone score, and the number of news, respectively. $TONEComove$ and $NNEWSComove$ are the news commonality measures as in Dang et al. (2015), defined as the R^2 obtained from the regression of a firm's news scores on the market's news scores. Panel A reports the contemporaneous relation between return comovement and partial correlation in news and Panel B reports the lead-lag relation. All control variables are defined in Equation (3.14). The table reports the regression results across firm-months with industry and month fixed effects. T-statistics based on industry-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $RetComove_{k,t}$	(2) $RetComove_{k,t}$	(3) $RetComove_{k,t}$	(4) $RetComove_{k,t}$
Panel A: Contemporaneous relation between return comovement and news partial correlation				
$LPCORR\_TONE_{k,t}$	0.0875***	0.0849***		
	(3.08)	(3.04)		
$TONEComove_{k,t}$		0.0117***		
		(4.88)		
$LPCORR\_NNEWS_{k,t}$			0.1188***	0.1119***
			(4.68)	(4.44)
$NNEWSComove_{k,t}$				0.0115***
				(5.91)
$LPCORR\_ROA_{k,t}$	0.1347***	0.1343***	0.1345***	0.1337***
	(4.75)	(4.77)	(4.75)	(4.69)
$Size_{k.t-1}$	0.1068***	0.1100***	0.1090***	0.1106***
	(4.56)	(4.77)	(4.57)	(4.63)
$BM_{k.t-1}$	0.0732***	0.0741***	0.0731***	0.0734***
	(2.99)	(3.05)	(3.00)	(3.02)
$Price_{k.t-1}$	0.0390	0.0384	0.0373	0.0365
	(1.06)	(1.05)	(1.00)	(0.98)
$R_{k.t-1}$	-0.1269*	-0.1284*	-0.1273*	-0.1269*
	(-1.80)	(-1.82)	(-1.78)	(-1.78)
$R_{k,t-2,t-7}$	0.1115***	0.1110***	0.1094***	0.1094***
	(8.54)	(8.57)	(8.68)	(8.63)
$RetStd_{k.t-1}$	-7.0902***	-7.0334***	-7.0335***	-7.0023***
	(-5.98)	(-5.95)	(-5.92)	(-5.91)
$Liquidity_{k,t-1}$	-0.0243***	-0.0242***	-0.0241***	-0.0240***
	(-3.97)	(-3.94)	(-4.00)	(-3.99)
$Analyst_{k,t-1}$	0.1567***	0.1559***	0.1488***	0.1473***
	(3.42)	(3.41)	(3.24)	(3.20)
$IO_{k,t-1}$	0.0994***	0.0992***	0.0969***	0.0966***
	(6.31)	(6.31)	(6.27)	(6.26)
Constant	-1.9635***	-1.9455***	-1.9030***	-1.8773***
	(-7.73)	(-7.60)	(-7.44)	(-7.29)
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	108,745	108,742	108,133	108,131
R-squared	0.4128	0.4133	0.4113	0.4117

Panel B: Lead-lag relation between return comovement and news partial correlation				
$LPCORR\_TONE_{k,t-1}$	0.0201***	0.0199***		
	(3.35)	(3.36)		
$TONEComove_{k,t-1}$		0.0005		
		(0.90)		
$LPCORR\_NNEWS_{k,t-1}$			0.0307***	0.0295***
			(5.80)	(5.86)
$NNEWSComove_{k,t-1}$				0.0019
				(1.48)
$LPCORR\_ROA_{k,t-1}$	0.0297***	0.0297***	0.0294***	0.0293***
	(5.29)	(5.30)	(5.02)	(4.94)
$RetComove_{k,t-1}$	0.7758***	0.7758***	0.7763***	0.7762***
	(58.22)	(58.29)	(57.86)	(57.69)
$Controls_{k,t-1}$	Yes	Yes	Yes	Yes
Industry fixed-effects	Yes	Yes	Yes	Yes
Month fixed-effects	Yes	Yes	Yes	Yes
Observations	108,433	108,429	107,817	107,814
R-squared	0.7664	0.7664	0.7663	0.7663

3.8 Conclusion

Developed from Veldkamp's (2006) competitive information market model, this chapter investigates how information is disseminated within an industry and its implications for stock return comovement. Empirical results reveal unidirectional information spillover from industry bellwether firms to their peer firms. News about bellwether firms exhibits significant influence on peer firms' stock prices, trading activity and analyst forecasts. Non-bellwether peers do not exhibit these relations. Further tests indicate that the intra-industry news spillover varies with industry competition, market states, news tone and industry peers' information environment. Using a firm's partial correlation in news (with other firms in the same industry) to gauge the firm's contribution in explaining news of other firms, we find that bellwether firms exhibit a higher partial correlation in news. Motivated by Veldkamp's (2006) prediction that information production can generate return comovement, we examine how news partial correlation is associated with return comovement. Results show that firms with more contributing news are associated with higher return comovement. Furthermore, firms with more contributing news also exhibit a lower degree of mispricing. Our findings provide supportive evidence for the positive relation between return comovement and price informativeness.

This chapter makes three main contributions to the literature. First, the finding on the news spillover from bellwether firms to industry peers validates Veldkamp's (2006) theoretical prediction that investors price assets using a common subset of information. Second, it contributes to the growing body of literature on the impact of media in financial market by showing that news production affects return comovement. Third, it adds to the long-standing debate on the information implication of return comovement by explicitly examining how the stock's informativeness is associated with return comovement and price efficiency.

Chapter 4

Style investing, investor attention and return predictability

4.1 Introduction

Investors choose among thousands of financial assets when allocating capital. To simplify asset allocation decisions, investors categorize assets into broad classes and allocate funds across various asset classes rather than at the individual asset level. There is growing evidence that investors often group stocks into categories based on shared commonalities. For example, existing studies document excess return comovement among stocks in the same index, with similar prices or dividend payout policies (Barberis et al., 2005; Boyer, 2011; Greenwood, 2007; Green and Hwang,2009; Hameed and Xie, 2019). Barberis and Shleifer (2003) refer to the categories investors use in making portfolio allocation decisions as "styles", and the process of allocating funds among styles as "style investing". In their study, the authors present a model in which investors allocate capital based on the relative performance of styles and predict that style investing generates momentum and reversals in style and individual asset returns, as well as comovement between individual assets and their styles.

Empirical research provides ample evidence of excess price comovement among different categorical stocks and style-level price predictability.[49] Prices are the outcome of information acquisition and trading activities. However, the extant research on style investing has largely focused on the outcome (i.e., excess return comovement and time-series return patterns) without necessarily investigating the information flows leading to these outcomes. This chapter aims to explore the asset pricing consequences of style investing from the information flow channel. Specifically, we study how investors allocate attention across different investment styles and whether the concentrated demand for information contributes to the style-related price patterns.

This chapter sits at the intersection of investor attention and style investing literatures. As discussed in Chapter 2, investor attention literature is based on the premise that investors have limited attention. As a result, they are more likely to consider securities that attract their attention. Empirical evidence indicates that investor attention is associated with significant capital market consequences (Da et al.,

[49] Categories identified by those studies include stock price (Green and Hwang, 2009), industry membership (Kallberg and Paquariello, 2008), stock index membership (Barberis et al., 2005), investment banking networks (Grullon, Underwood, and Weston, 2014), and analyst following (Muslu, Rebello, and Xu, 2014). Examples of studies on style-level return predictability include Froot and Teo (2008), Jame and Tong (2014), Kumar (2009), Teo and Woo (2004) and Wahal and Yavuz (2013).

2011; Fang and Peress, 2009; Hirshleifer and Teoh, 2003; Li and Yu, 2012). The style investing literature argues that investors categorize assets into styles and chase styles that have performed relatively well in the past. The concentrated shift in demand creates predictability in asset returns. Our study links the two literatures by examining information flows (captured by investor attention) underlying style investing and its implication for return predictability. Our research objectives are twofold. First, we investigate what drives cross-style attention allocation. Second, we examine whether attention comovement helps predict within-style return comovement, and whether style-level attention helps explain the variation in cross-sectional stock returns.

Barberis and Shleifer (2003) predict that many investors allocate funds at the style level (style switchers), constantly shifting capital from loser styles to winner styles. We therefore expect that style portfolios with more extreme past performance would attract more investor attention. Furthermore, if investors systematically allocate attention to stocks with similar style, their attention should comove. As a result, the capital allocation may similarly comove, generating comovement in returns within the style. For the same reason, when a stock is reclassified into a new style, investors would treat it as a part of the new investment category. Consequently, its comovement with that style rises, and comovement with the old style falls.

In the other strand of research, Barber and Odean (2008) suggest that attention leads to net-buying behaviour, which causes higher returns over the short run and price reversals over the long run. Linking Barber and Odean's (2008) price pressure hypothesis to style investments, if investors pay attention to a certain style, they may push up prices of the stocks in this style away from their fundamentals, causing high returns over a short period and subsequent price reversal. The consequence from changes in attention would result in positive autocorrelation in style returns in the short run and reversals in the long run. This chapter aims to test these predictions.

To capture the amount of attention paid to a specific firm, we employ two measures that have been used in prior research – the number of analyst forecast revisions and news coverage. Previous studies suggest that earnings forecasts issued by financial analysts can significantly affect institutional investors' investment decisions (Barber, Lehavy, McNichols, and Trueman, 2006; Ding, Chen, and Wu, 2014). Also, news media plays an important role in disseminating information to a broad audience, especially to individual investors (Fang and Peress, 2009). Barber and Odean (2008) show that stocks in the news tend to grab individual investors' attention. Therefore, the two measures aim to capture attention from different investor groups. Analyst forecast revisions capture attention from more sophisticated institutional investors, while news coverage captures attention from retail investors. As this chapter focuses on attention to different investment styles, for this purpose, we identify styles using the size and value-growth grids following the extant literature.

We first investigate whether prior style performance plays a role in driving investor attention. Both style- and firm-level analyses suggest that investor attention is positively related to the absolute prior style returns. At the style-level, styles that demonstrate extreme performance over the past one, three and six months are associated with higher attention. Prior literature documents a positive relation between abnormal stock returns and firm-specific attention. If, as in Barberis and Shleifer (2003), investors invest at the style level rather than at the individual asset level, the association between style returns and firm-specific attention should persist after controlling for individual stock returns. To test this, we regress firm-level attention on prior style returns while controlling for individual stock returns. The results show that firm-level attention is positively associated with style returns, but its relation with the firm's past returns is weak. This finding is consistent with Barberis and Shleifer's (2003) prediction of style investing and style chasing behaviour.

We then examine the implication of investor attention for return comovement within the style. Following a similar approach to previous chapters, we develop an R^2 measure of attention comovement to identify the amount of attention a firm receives that is explained by the attention paid to its style. Specifically, we regress firm-specific attention on style-level attention, which is created by aggregating firm-level attention within each style. The resulting R^2 is used as a proxy for attention comovement, with a higher value of R^2 corresponds to stronger attention comovement. We investigate whether a stock's attention comovement helps explain its return comovement with the style.

In general, the results suggest that when investors systematically seek out information for a similar style of stocks, their attention comoves. This leads to comovement in trading activity and stock returns. The style-level analysis shows that, on average, high attention-grabbing styles demonstrate stronger attention comovement, return comovement and trading comovement. Within each style, attention comovement has predictive power for return comovement as well as trading comovement even after controlling for comovement in fundamentals. This suggests that comovement in returns and trading is partially driven by the actions of investors who view individual firms in the context of categories such as styles.

To further support this finding, we investigate the changes in attention comovement, return comovement and trading comovement following a stock's style reclassification. After a stock enters a new style, its attention comovement with the new (old) style rises (falls). Similar patterns are also observed for return comovement and trading comovement. These findings add additional evidence for Barberis and Shleifer's (2003) style investing model and provide information flow explanation for excess return comovement within the style.

Finally, we investigate whether style-level attention helps explain the variation in cross-sectional style returns and stock returns. Consistent with Barber and Odean's (2008) price pressure hypothesis, we find that high style-level attention is associated with high contemporaneous style returns and return reversals over the long run. For example, using the number of analyst forecast revisions as the attention measure, the top attention style quintile outperforms the bottom attention style quintile by 46 bps in the concurrent month, but underperforms by 16 bps over the subsequent 12 months. To test Barberis and Shleifer's (2003) prediction of autocorrelation in style returns, we examine the impact of style attention on future style returns conditional on prior style performance. The results indicate that style chasing temporarily pushes price away from fundamentals, generating momentum in the short run and price reversals over the long run. High attention is associated with high returns in the short-term and low returns in the long-term for prior outperforming styles, and the signs are opposite for prior underperforming styles.

We also examine the association between style-level attention and expected individual stock returns. The existing literature has established strong evidence that the level of attention a specific firm receives has a significant impact on its stock returns. To distinguish the effect of style-level attention from firm-specific attention, we first sort stocks into quintiles based on firm-specific attention. Within each attention group, we examine how an individual stock's returns are affected by attention paid to the style where the firm belongs. The results show that short-term price increase and long-term reversal patterns are prevalent among all firm-attention groups, suggesting that style-level attention provides incremental prediction for future stock returns.

In summary, this chapter contributes to the literature by linking style investing with investor attention. First, the style investing literature assumes that investors categorize stocks into different styles and chase styles based on their relative performance. We explicitly test this assumption by investigating how style performance affects investors' attention allocation across styles. Our results provide empirical evidence for the underlying assumptions of Barberis and Shleifer's (2003) theoretical work. Second, standard asset pricing models cannot fully explain the empirically documented style-related return predictability, which calls for an exploration of behavioural explanations. We contribute to the literature by showing that investor attention helps explain the excess within-style return comovement and autocorrelation in style returns. Third, existing attention literature has largely focused on the level of attention that a specific firm receives and its asset-pricing implication. Drake et al. (2017) argue that due to social interaction and categorial thinking, investor attention is likely a social construct (in addition to being an individual construct). Our finding on the style-level attention comovement provides support for this view. Furthermore, we show that comovement of investor attention has market consequences in that it is positively associated with excess return comovement within the style.

The remainder of this chapter is structured as follows. Section 4.2 reviews the literature on style investing and investor attention. Section 4.3 summarises our research questions and presents hypotheses. Section 4.4 describes the data sources and methodology. Sections 4.5 and 4.6 present the empirical results and robustness checks, respectively. Section 4.7 concludes the chapter.

4.2 Literature review

4.2.1 Style investing

One of the clearest mechanisms of human thoughts is classification, the grouping of objects into categories based on some similarity among them (Rosch and Lloyd, 1978; Wilson and Keil, 2001). Classification of large numbers of objects into categories is also pervasive in financial markets. Barberis and Shleifer (2003) argue that, to simplify portfolio decisions, many investors first categorize assets into broad classes such as small-cap stocks, value stocks, oil industry stocks, or government bonds, and then allocate funds across different asset classes rather than at the individual asset level. The asset classes investors use for making asset allocation decisions are known as styles, and the process of allocating funds among styles is known as style investing.

Barberis and Shleifer (2003) analyse financial markets in which many investors pursue style investing. Their model generates a rich set of predictions, some of which have received empirical support. The model is built on the assumption that many investors allocate funds based on their past performance, moving into styles that have performed well in the past and withdrawing funds from styles that have performed poorly. This assumption is validated by subsequent empirical studies. For example, Kumar (2009) show that individual investors systematically shift their preferences across style portfolios (small vs. large, value vs. growth) based on the past style returns and earnings. Jame and Tong (2014) find that retail investors tend to follow industries that have performed well over the past two years.

Barberis and Shleifer (2003) provide two interesting and empirically testable predictions on the implication of style investing for asset prices. First, the prices of assets within the same style will comove more than their comovement in fundamentals, while the prices of assets in different styles will comove less than their comovement in fundamentals. Second, style investing generates momentum and reversals in both style and individual asset returns. There is a growing list of empirical studies that provide supportive evidence for these predictions.

Froot and Dabora (1999) find that prices of 'Siamese twin' stocks traded on different exchanges (e.g., Royal Dutch and Shell) do not move in lockstep, but rather are correlated with the movements of their respective exchanges. Pirinsky and Wang (2006) document that stocks of companies that change their headquarters location experience a decrease in their comovement with stocks from the old location and an increase in their comovement with stocks from the new location. Lee et al. (1991) show that discounts of closed-end funds that are listed on the same exchange but hold different securities move together. Barberis et al. (2005) find that stocks added to the S&P 500 index begin to covary more with other members of the index, and covary less with stocks outside the index. Greenwood (2007) provides similar evidence for the Nikkei 225. Claessens and Yafeh (2012) provide cross-country evidence on comovement of newly added stocks with national market indices. Using membership switching between S&P/Barra Value and Growth indices as a setting, Boyer (2011) shows that stocks begin to covary more with the index they join and less with the index they leave, even though these labels sometimes have little connection with underlying fundamentals.[50]

In addition to the aforementioned studies which reveal excess comovement among assets in the same market or index, a large body of literature documents common factors among stocks with similar characteristics. Several studies provide evidence that investors tend to categorize stocks based on size and book-to-market (e.g., Fama and French, 1995; Kumar, 2009; Teo and Woo, 2004; Wahal and Yavuz, 2013). Some studies show that industry is a common source of systematic variance that generates price comovement among firms (Brandt, Brav, Graham, and Kumar, 2009; Campbell et al., 2001; Irvine and Pontiff, 2008; Kallberg and Pasquariello, 2008; Jame and Tong, 2014). Green and Hwang (2009) document that stocks which undergo stock splits experience an increase in comovement with low-priced stocks and a decrease in comovement with high-priced stocks. Hameed and Xie (2019) find that stocks that initiate dividends tend to comove more with other dividend-paying stocks and comove less with non-dividend payers. Kumar, Page, and Spalt (2016) show that lottery-like stocks comove strongly with one another. Muslu et al. (2014) and Israelsen (2016) show that stocks with similar sets of analysts exhibit more excess comovement. Grullon et al. (2014) and Anton and Polk (2014) show that a shared investment bank network and mutual fund ownership can lead to excess returns comovement. These results are consistent with the prediction that investors categorize stocks into different styles, and the style investing generates excess comovement of assets within styles.

The prediction that style investing generates momentum and reversals in style and individual asset returns is also supported by empirical evidence. Earlier studies show that style-level momentum and value strategies are profitable. Moskowitz and

[50] S&P/Barra define their Value and Growth indices by dividing all S&P 500 stocks into two mutually exclusive categories according to simple mechanical rules. Stocks in the S&P 500 with a book-to-market ratio above a given boundary constitute the Value index and all others make up the Growth index. The boundary is reset and the indices are rebalanced every June and December so the two indices have equal market cap. In doing so, some stocks are reclassified from Value to Growth even after their book-to-market ratios have risen, and vice versa.

Grinblatt (1999) and Asness, Liew, and Stevens (1997) successfully apply momentum strategies to industry portfolios and country portfolios, respectively. Lewellen (2002) finds that momentum strategies based on size and book-to-market portfolios are at least as profitable as individual stock momentum. Kumar (2009) finds that retail investors herd into similar size and book-to-market styles and finds evidence of style-level momentum. Chen and De Bondt (2004) document style momentum within the S&P 500 index. Asness et al. (1997) show that a value strategy works well with country portfolios. Following Morningstar style classification system, Teo and Woo (2004) categorize stocks into small, mid-cap, or large, and growth, blend, or value. They document style-level momentum at quarterly horizons and reversals at annual horizons. More recent literature provides evidence that style investing also plays a role in the predictability of individual asset returns. Wahal and Yavuz (2013) identify styles using size and value-growth grids and show that style returns are significant predictors of future returns. Jame and Tong (2014) show that industry-based investment styles generate short-term momentum and longer horizon reversals in both style and individual stock returns.

4.2.2 Investor attention and return predictability

This chapter also draws on the investor attention literature. This strand of literature is motivated by psychological evidence that attention is a scarce cognitive resource. Since investors cannot pay attention to all stocks, they must be selective about the particular stocks they choose to follow (Hirshleifer and Teoh, 2003). As discussed in Chapter 2, there has been a growing body of literature investigating the level of investor attention a firm receives and the impact on its stock prices. Those studies suggest that investor attention has significant impact on stock prices, and high attention is associated with high returns over the short run and price reversals over the long run (Barber and Odean, 2008; Merton, 1987).

Prior studies on investor attention have largely focused on the level of attention that a specific firm receives and its asset-pricing implication. Most recent literature suggests that, as a result of social forces and interactions, investors collectively focus on similar firms and therefore their attention comoves.[51] Thus, attention is not just a firm-level construct, it is a macro construct as well. This leads to the view that correlated information flows help explain the excess return comovement among similar categorial stocks (Drake et al.,2017).

[51] Shiller (1989) calls investing a "social activity" in which investors discuss, read, and gossip about investment.

Mondria (2010) provides a theoretical framework to study the impact of investor attention allocation on price comovement. The model assumes that instead of processing information on individual assets, investors with limited information processing capacity choose to observe one linear combination of asset payoffs as a private signal. When investors use this private signal to update information about multiple assets, changes in one asset affect other assets' prices, leading to price comovement. Peng and Xiong (2006) study the learning process of investors who are subject to attention constraints and behavioural biases in information processing. They show that limited investor attention leads to category-learning behaviour. The interaction of the category-learning behaviour with investor overconfidence generates excess return comovement.

The prediction on the effect of correlated attention on return comovement is supported by several recent empirical findings. Motivated by the style investing literature (Barberis et al., 2005), Drake et al. (2017) show that the amount of attention a firm receives comoves with the amount of attention paid to its industry and the market as a whole. Furthermore, the study shows that comovement in attention is positively associated with comovement in stock returns. Using cosearch data for stocks from Yahoo! Finance, Leung, Agarwal, Konana, and Kumar (2016) analyse users' aggregate search correlations. The study identifies 50-79 search clusters at different time points representing 230-349 stocks. Leung et al. (2016) find that the stock returns within the search clusters are strongly correlated. When a stock enters (departs) a cluster, the focal stock return comoves (detaches) with the cluster returns. Those findings provide evidence for the information flows underlies return comovement.

4.3 Research questions and statement of hypotheses

4.3.1 Research questions

Motivated by the existing literature, this chapter aims to investigate the following research questions: (1) What drives cross-style attention allocation? (2) How does cross-style attention allocation affect stock returns?

4.3.2 Hypotheses

Empirical studies provide evidence that investors tend to invest in styles based on their past performance, switching funds from underperforming styles to outperforming styles. If investors pay attention to style performance, investor attention is expected to be clustered in styles with extreme past performance. Accordingly, our first hypothesis is:

H1: Investor attention is positively related to the absolute past style returns.

Barberis and Shleifer (2003) predict that investors use categories (styles) to simplify their asset allocation decisions. When investors systematically seek out information for similar categorical stocks or experience correlated shocks to the demand for information, attention to stocks within the same style comoves. It is plausible that investors are more likely to systematically search for information about styles that attract their attention. Therefore, we expect that attention-grabbing styles exhibit stronger attention comovement. Accordingly, our second hypothesis is:

H2: Attention comovement is stronger among attention-grabbing styles.

Prior studies document excess return comovement among stocks within the same style. The return comovement cannot be fully explained by common factors in cash flows or risks, which promotes the investigation of behavioural explanations. If investor attention comoves with the broader level attention paid to its style, it potentially explains the excess comovement in stock returns documented in previous research. To investigate this conjecture, our third hypotheses are:

H3a: There is a positive relation between attention comovement and return comovement within the style.

Building on the same argument, we further assert that when a stock is reclassified into a new style, its attention comoves more (less) with the new (old) style after reclassification. As a consequence, the stock's return comovement with the new (old) style increases (decreases).

H3b: A stock's attention comoves more with the new style and comoves less with the old style after style reclassification.

H3c: A stock's returns comove more with the new style and comove less with the old style after style reclassification.

Both theoretical and empirical evidence suggests that high investor attention leads to a contemporaneous rise in stock price and a subsequent price reversal. If investors categorize stocks into groups when making investment decisions, style-level attention should exert significant impact on both style and individual stock returns. Accordingly, our forth hypotheses are:

H4a: Style-level attention is associated with a short-term increase and long-term reversal in style returns.

H4b: Style-level attention is associated with a short-term increase and long-term reversal in individual stocks returns within the style.

4.4 Data and methodology

4.4.1 Data

We consider all stocks with shares codes of 10 and 11 trading on the NYSE, Amex, and Nasdaq in the Center for Research in Security Prices (CRSP) database. Accounting data are obtained from Compustat. We use analyst forecast revisions and news coverage to capture investor attention. The availability of news data restricts the sample to the period between January 2003 and March 2016.

Our first measure of attention, analyst forecast revision ($Analyst_i$), is the number of earnings forecast revisions made for a given firm from the Institutional Brokers' Estimate System (I/B/E/S) detail history file. Given that our goal is to capture the attention from sophisticated investors like financial analysts and institutional investors, we do not restrict our analysis to any single type of forecast revision (e.g., 1-year-ahead earnings forecasts). Following Frankel, Kothari, and Weber (2006), we count the number of unique earnings per share revisions issued for each firm, including forecasts for all time horizons (i.e., quarterly, one-year, two-year, and all other horizons).[52]

The second measure of attention is news coverage, $News_i$, which is equal to the number of articles issued by the business press for a firm. We use news data collected from the Thomson Reuters News Analytics (TRNA), which is available for the period January 2003 - March 2016. TRNA sources and analyses firm-level news from major news outlets such as Dow Jones Newswires, the Wall St Journal, Reuters, and local newspapers. It not only provides the news coverage related to a firm, but also a tone score and a relevance score for each news item.[53] Previous studies suggest that stocks in the news are more likely attention-grabbing stocks for individual investors (Fang and Peress, 2009; Barber and Odean, 2008). Therefore, we use news coverage to capture retail investors' attention.

[52] We count all revisions on the same date by the same analyst for the same firm as a single analyst revision.

[53] TRNA also provides a score of how relevant a news item is to a given firm. If a story mentions multiple firms in the contents, the firm with the most mentions is given highest relevance score. According to Thomson Reuters, the relevance score allows the distinction between the three cases: (1) when the relevance score is greater than 0.8, the company is "one of the determinant players" in the article; (2) when the relevance score is between 0.8 and 0.2, the firm is "one of several mentioned substantively in the article"; and (3) when the relevance score is less than 0.2, the company is a "minor player" in the article. To ensure that the news article is highly relevant to the firm, we require the relevance score to be greater than 0.8 as a robustness check. Results are very similar.

4.4.2 Main variables

4.4.2.1 Measures of attention shock

We estimate attention shock ($AttnShock$) for each firm at a monthly frequency. Using analyst forecast revisions as the attention measure, attention shock for firm i in month t is defined as:

$$AnalystShock_{i,t} = \log\left(1 + Analyst_{i,t}\right) - \log\left(1 + AVGAnalyst_{i,t-12,t-1}\right) \tag{4.1a}$$

where $Analyst_{i,t}$ is the number of analysts forecast revisions made for firm i during month t. $AVGAnalyst_{i,t-12,t-1}$ is the average number of analyst forecast revisions over the past 12 months. The average $Analyst$ over a 12-month window not only captures the normal level of attention, but also removes the seasonality in the data. A large positive $Analyst\_shock$ represents a surge in attention and can be compared in the cross-section of stocks.

Similarly, when news coverage is used to proxy for investor attention, attention shock is defined as:

$$NewsShock_{i,t} = \log\left(1 + News_{i,t}\right) - \log\left(1 + AVGNews_{i,t-12,t-1}\right) \tag{4.1b}$$

where $News_{i,t}$ is the number of news articles on firm i in month t, and $AVGNews_{i,t-12,t-1}$ is the monthly average news coverage over the past 12 months.

4.4.2.2 Measures of comovement

Our analyses require the construction of attention comovement, return comovement and turnover comovement measures. Following the method used in measuring return comovement, we regress firm-specific attention on the style-level attention using weekly data from July of year t-1 to June of t:[54]

$$Attention_{i,w} = \alpha_i + \beta_1 Attention_{Style,w} + \varepsilon_{i,w} \tag{4.2}$$

where i denotes firms and w denotes weeks. $Attention$ represents the attention measure using the number of analyst forecast revisions ($Analyst$) or news coverage ($News$). $Attention_{Style}$ is the style-level attention measure, computed as the equal-weighted attention for the style for a given week (excluding firm i). We estimate Equation (4.2) for each firm and each period to obtain R^2. Attention comovement is estimated by taking the log transformation of the R^2:

$$AttnComove_{i,t} = \ln\left(\frac{R^2}{1-R^2}\right) \tag{4.3}$$

Using this approach, we estimate an attention comovement variable for each of the attention measures and denote it as $AnalystComove$ and $NewsComove$, respectively. A high $AttentionComove$ value indicates that a firm's attention is more closely tied to its style-level attention.

Consistent with the estimation of attention comovement, the comovement in stock returns for each firm in each period is estimated using weekly stock returns from July of year t-1 to June of t, which is measured as the R^2 from the following regression:

$$Ret_{i,w} = \alpha_i + \beta_1 Ret_{Style,w} + \varepsilon_{i,w} \tag{4.4}$$

where $Ret_{i,w}$ is the weekly return for firm i in week w. $Ret_{Style,w}$ is the equal-weighted style return in week w (excluding firm i). To obtain an estimate of return comovement, we take the log transformation of the regression R^2 from Equation (4.4):

$$RetComove_{i,t} = \ln\left(\frac{R^2}{1-R^2}\right) \tag{4.5}$$

[54] Different from $AttnShock$ variable which is measured at monthly frequency, we use weekly $attention$ data when constructing annual attention comovement measure. This is to make sure there are sufficient number of observations for the regression analysis to have adequate statistical power. In addition, since the style portfolios are formed at the end of June each year, we use the 12-month period from July of year t-1 to June of t as the estimation window.

An analogous procedure is employed to estimate comovement in stock turnover. Specifically, turnover comovement is measured as the R^2 from the regression of individual stock's turnover on the style-level turnover using weekly data from July of year *t*-1 to June of year *t*:

$$Turnover_{i,w} = \alpha_i + \beta_1 Turnover_{Style,w} + \varepsilon_{i,w} \tag{4.6}$$

where $Turnover_{i,w}$ is stock turnover for stock *i* in week w, calculated as the sum of daily turnover across the week. Daily turnover is defined as daily trading volume divided by the share outstanding. $Turnover_{Style,w}$ is the equal-weighted stock turnover for a week w (excluding firm i). Stock turnover comovement is estimated by taking the log transformation of the R^2 from Equation (4.6)

$$TOComove_{i,t} = \ln\left(\frac{R^2}{1-R^2}\right) \tag{4.7}$$

4.4.3 Methodology

4.4.3.1 Style identification

We use size and book-to-market ratios (i.e., value-growth orientations) to identify investment styles in this study. There are several reasons for choosing this style identification. First, empirical evidence suggests these two firm characteristics are important in asset returns and they are widely used by managed funds (Boyer, 2011; Cooper, Gulen, and Rau ,2005; Froot and Teo, 2008; Kumar,2009). Second, we need styles that are mutually exclusive, and style portfolios generated by size and book-to-market ratios fit this criteria. By construction, a stock classified in a particular style cannot be placed in an alternative style portfolio in the same period.

We calculate size and book-to-market as in Fama and French (1992). At the end of June each year, we split stocks into quintiles based on the market value of equity. Independently, we split stocks into quintiles based on book-to-market, which is the book equity for the fiscal year ending in year *t*-1 divided by the market value at the end of December of year *t*-1. We use NYSE breakpoints for size, which are downloaded from Professor Kenneth French's website.[55] For book-to-market breakpoints, we use the full set of securities. The two independent sorts result in 5×5 size and book-to-market style

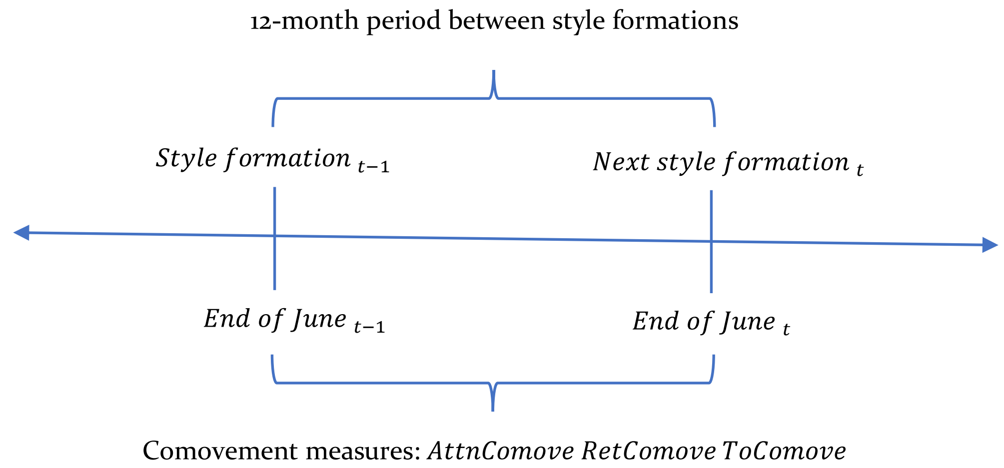
[55] NYSE size breakpoints better proxy for the styles followed by investors, they correlate more closely with the classifications employed by investors, such as those constructed by Russell, S&P Barra, and other such organizations.

portfolios at the end June each year. For each style portfolio, we calculate monthly equal-weighted portfolio returns over the following 12 months.[56]

Figure 4.1 provides a visual portrayal of the style formation timeframe and measurements of comovement variables. As demonstrated, a stock is assigned to a style portfolio at the end of June of year t-1. At the end of June of year t, comovement variables ($AttnComove$, $RetComove$ and $ToComove$) for this stock are estimated from a regression of the firm-specific variable on the equal-weighted style variable using weekly data from July of year t-1 to June of year t (see Section 4.4.2.2).

Figure 4.1 Style formation timeline and comovement variable measurement

Figure 4.1 presents the timeframe for constructing the within-style comovement variables. For each stock i, it is assigned to a style portfolio at the end of June of year t-1. At the end of June of year t, its comovement variables ($AttnComove$, $RetComove$ and $ToComove$) are derived from a regression of the firm-specific variable on the equal-weighted style variable (excluding firm i), using the weekly data over the 12-month period from July of year t-1 to June of year t. We then log transform the R^2 of the regression to measure the attention, return, or turnover comovement.



[56] We calculate the value-weighted portfolio returns as a robustness check, results are qualitatively similar.

4.4.3.2 Style performance and investor attention

Our analysis begins with the examination of Barberis and Shleifer's (2003) style chasing hypothesis. We investigate whether investors allocate attention across different styles based on their past performance by estimating the following regression across styles and months:

$$\begin{aligned} AttnShock_{s,t} &= \alpha_0 + \beta_1 Size_{s,t} + \beta_2 BM_{s,t} + \beta_3 \left|Ret_{s,t-1}\right| + \beta_4 \left|Ret_{s,t-3,t-2}\right| \\ &\quad + \beta_5 \left|Ret_{s,t-6,t-4}\right| + \beta_6 \left|Ret_{s,t-12,t-7}\right| + \varepsilon_{s,t} \end{aligned} \tag{4.8}$$

where the dependent variable $AttnShock_{s,t}$ is the attention shock for style s in month t, calculated as the average attention shock, measured by $AnalystShock$ or $NewsShock$, across all stocks within the style. Independent variables include $Size_{s,t}$ and $BM_{s,t}$, which are equal to the style average market capitalization and the style average book-to-market ratio in month t (both in natural logs). The main variables of interest are measures of absolute past style returns, ranging from the prior one month return to the return over the prior 7 to 12 months. The regression allows us to explore how investor attention is affected by past style returns over different horizons. If investors pay more attention to extreme style portfolios as suggested by the literature, we expect coefficients on the prior style returns to be significantly positive.

Furthermore, if investors invest at the style level rather than at the individual asset level, we expect a firm's attention is related to its style returns after controlling for the firm's past returns. Accordingly, we perform the following regression across firms and months:

$$\begin{aligned} AttnShock_{i,t} &= \alpha_0 + \beta_1 Size_{i,t} + \beta_2 BM_{i,t} + \beta_3 \left|Ret_{s,t-1}\right| + \beta_4 \left|Ret_{s,t-3,t-2}\right| \\ &\quad + \beta_5 \left|Ret_{s,t-6,t-4}\right| + \beta_6 \left|Ret_{s,t-12,t-7}\right| + \beta_7 \left|Ret_{i,t-1}\right| \\ &\quad + \beta_8 \left|Ret_{i,t-3,t-2}\right| + \beta_9 \left|Ret_{i,t-6,t-4}\right| + \beta_{10} \left|Ret_{i,t-12,t-7}\right| + \varepsilon_{i,t} \end{aligned} \tag{4.9}$$

where $AttnShock_{i,t}$ is the attention shock for firm *i* in month t. $Size_{i,t}$ and $BM_{i,t}$ are the logarithm of firm *i*'s market capitalization and book-to-market, respectively. $|Ret_s|$ represents the absolute return of stock *i*'s style and $|Ret_i|$ is stock *i's* absolute return. If style-level performance plays an incremental role in explaining firm-specific attention, coefficients on $|Ret_s|$ are expected to remain significantly positive after controlling for firm-level returns.

4.4.3.3 Attention comovement and return comovement

Barberis and Shleifer (2003) predict that style investing generates excess comovement in assets within the same style. We explore an information flow channel for the style-wide price comovement by investigating how attention comovement is associated with return comovement and turnover comovement within the style. To do so, we estimate the following panel regression:

$$\begin{aligned} RetComove_{i,t}\ (or\ TOComove_{i,t}) &= \alpha_0 + \psi_1 AttnComove_{i,t} + \psi_2 ROA_{i,t} + \psi_3 Size_{i,t} + \psi_4 BM_{i,t} \\ &+ \psi_5 SalesGrowth_{i,t} + \psi_6 IO_{i,t} + \psi_7 Analyst_{i,t} + \psi_8 StdROA_{i,t} \\ &+ \psi_9 AbsRet_{i,t} + \psi_{10} Turnover_{i,t} + \psi_{11} Price_{i,t} \\ &+ \psi_{12} ROACom_{i,t} + \mathcal{E}_{i,t} \end{aligned} \quad (4.10)$$

where $RetComove_{i,t}$ and $TOComove_{i,t}$ are stock *i*'s return comovement and turnover comovement with its style over the period from July of year *t-1* to June of year *t*, respectively. $AttnComove_{i,t}$ is stock *i*'s attention comovement with the style, measured by $AnalystComove$ or $NewsComove$. We also control for several firm characteristics that are potential determinants of return and turnover comovement identified by prior studies. Those variables include the firm's market capitalization (*Size*), book-to-market ratio (*BM*), stock price (*Price*), the fraction of shares outstanding held by institutional investors (*IO*), the number of analysts following *(Analyst*), the absolute buy-and-hold monthly return (*AbsRet*), and the stock turnover (*Turnover*). *Size, BM, Price, IO, and Analyst* are natural logged. In addition, to control for factors that are related to firm fundamentals, we include return on assets (*ROA*), sales growth (*SalesGrowth*), standard deviation of return on assets (*StdROA*), and comovement in return on assets (*ROACom*). See Appendix A4.1 for more details of these specific variables.

The primary coefficient of interest is ψ_1, which captures the association between return comovement (turnover comovement) and attention comovement. Hypothesis 3a predicts that correlated attention leads to correlated trading activity and price movement. Therefore, ψ_1 is expected to be positive.

4.4.3.4 Changes in comovement after style reclassification

Behavioural theories suggest that, due to category investment behaviour, assets in the same style comove too much and assets in different styles comove too little. Reclassifying an asset into a new style then raises its correlation with that style. To examine this prediction, we investigate how a stock's comovement with its old and new styles changes following style reclassification, defined as a stock migrating to a different assignment in the 5×5 grids.

We first examine how a stock's comovement in attention changes after style reclassification. In the spirit of Barberis et al. (2005), we estimate the following bivariate model for each event stock i for the period before the style reclassification and the period after the style reclassification, separately, and record the changes in the old style and new style betas, $\Delta\beta_{i,OldStyle}$ and $\Delta\beta_{i,NewStyle}$:

$$Attn_{i,t} = \alpha_i + \beta_{i,OldStyle}Attn_{OldStyle,t} + \beta_{i,NewStyle}Attn_{NewStyle,t} + \varepsilon_{i,t} \tag{4.11}$$

where $Attn_{i,t}$ is the level of investor attention of stock i in week t, measured by the number of analyst forecast revisions ($Analyst$) or news coverage ($News$). $Attn_{OldStyle,t}$ and $Attn_{NewStyle,t}$ are the equal-weighted attention of stock i's pre- and post-reclassification styles, respectively. To avoid spurious effects, we remove the contribution of the event stock i from $Attn_{OldStyle}$ when estimating the regression for the pre-reclassification period, and remove stock i from $Attn_{NewStyle}$ when estimating the regression for the post-reclassification period. For each event stock, we set an 11-month interval before the event month as the pre-event window, and 11 months after the event month as the post-event window. Since we rebalance style portfolios at the end of June every year, the pre-event window is from July of year t-1 to May of year t, and the post-event window is from July of year t to May of year t+1.[57]

We employ the same method to estimate the changes in return comovement and turnover comovement subsequent to style reclassification. For each event stock, we regress its returns (turnover) on the returns (turnover) of both old and new styles. We estimate the regressions over the 11-month pre-event window and the 11-month post-event window, separately, and observe the changes in the old style and new style betas. Hypotheses 3b and 3c predict that, after the style reclassification, a stock's comovement with the new (old) style raises (falls). Thus, $\Delta\beta_{i,OldStyle}$ is expected to be significantly negative, and $\Delta\beta_{i,NewStyle}$ is expected to be significantly positive.

[57] We use the the 11-month rather than 12-month pre- and post-event window to skip the portfolio formation month.

4.4.3.5 Style-level attention and cross-sectional returns

The final empirical test is on whether style-level attention contributes to the return predictability in style returns. We use portfolio sorts to conduct the investigation. At the end of each month, style portfolios are sorted into quintiles based on their level of attention shock. We then calculate returns for each attention quintile. To capture the dynamic relation between style attention and style returns, we consider a variety of portfolio formation and holding periods, including the concurrent month, 1 month – 1 month, 3 month – 3 month, 6 month – 6 month, and 12 month – 12 month strategies. For each strategy, style-level attention shock is calculated as the monthly average attention shock over the formation period. Hypothesis 4a predicts that style-level attention results in a short-term price increase and long-term reversal in style returns. Therefore, we expect the high attention styles to outperform the low attention styles in the short run, and underperform the low attention styles in the long run.

We also investigate whether style-level attention contributes to the variation in cross-sectional individual stock returns. In each month, we first sort stocks into quintiles based on firm-level attention. Within each firm-level attention quintile, we further sort stocks into quintiles based on style-level attention, resulting in 25 portfolios. Similar to the style-level analysis, we also consider different formation and holding periods. If style-level attention provides incremental explanation for the cross-sectional stock returns as predicted by Hypothesis 4b, we expect the association between style-level attention and return predictability to remain after controlling for firm-specific attention.

4.4.4 Summary statistics

Table 4.1 reports descriptive statistics for the key variables used in the analysis. Panel A presents the summary statistics for firm characteristics. On average, there are 6 forecast revisions made for each firm every month. The average number of news articles for firms in our sample is 17 per month. Generally, firms experience more negative attention shocks than positive attention shocks throughout our sample period, with a mean of -0.1810 for $AnalystShock$ and -0.1766 for $NewsShock$.

The mean $Analyst\ R^2$ is equal to 0.2222, ranging from a minimum value barely distinguishable from zero to a maximum value of 0.8595. This suggests that, on average, more than a fifth of the variation in firm-specific attention can be explained by the style-level attention. Also, we document an average News R^2 of 13.01%. These results are comparable to Drake et al. (2017) who examine the industry/market components of investor attention, and document a mean Analyst R^2 of 21.9% and a mean News R^2 of 14.5% over the period 2007-2011. Given that analyst revisions are more likely to capture institutional investor attention and news coverage captures retail investor attention, the difference between Analyst R^2 and News R^2 suggests that the within style attention comovement is stronger among more sophisticated investors than the retail investors. Since the attention R^2 measures are bounded between zero and one, we take a logarithmic transformation in our analysis, denoting them as $AnalystComove$ and $NewsComove$. The average $AnalystComove$ and $NewsComove$ are -1.9070 and -2.9850, respectively. Employing an analogous method, we document a mean value of -2.0078 for return comovement with the raw R^2 of 21.04%, and a mean value of -3.0515 for turnover comovement with the raw R^2 of 12.09%.

The mean market capitalization of firms in our sample is $3.6051 billion and the average book-to-market ratio is 0.7250. A typical firm in our sample is covered by about 7 analysts and has 53.32% firm share outstanding held by institutional investors. The average annual sales growth is 12.04%. The majority of firms in our sample are profitable with a median ROA of 1.71%.

Panel B displays the cross-sectional correlations among the key variables. Since size, book-to-market, and institutional ownership are logged in the regressions to follow, the correlations are also based on natural logs of these variables. $AnalystComove$ and $NewsComove$ has a correlation of 0.21. The relatively low correlation indicates the two measures capture attention from different investor groups. Overall, the correlations among the key variables are low. Unsurprisingly, size, institutional ownership and analyst coverage are closely correlated. Also, these three variables are positively correlated with return comovement and turnover comovement.

Table 4.1 Summary statistics

This table presents descriptive statistics for the key variables used in the study. Panel A reports summary statistics for firm characteristics. $Analyst$ is the number of analyst forecast revisions made for a firm in each month. $News$ is the number of news articles about a firm in each month. $AnalystShock$ and $NewsShock$ are attention shock measures. $AnalystShock$ ($NewsShock$) is calculated as the monthly $Analyst$ ($News$) minus the average $Analyst$ ($News$) over the past 12 months as described in Equation 4.1a (4.1b). $Analyst\ R^2$ ($News\ R^2$) is the attention R^2 estimated from regressing weekly firm-specific attention on the equal-weighted style-level attention over a 12-month window (excluding the firm), where attention is measured by $Analyst$ ($News$). Similarly, $Return\ R^2$ ($Turnover\ R^2$) is the R^2 estimated from the firm's weekly stock returns (turnover) regressed on the equal-weighted style returns (turnover) over a 12-month window (excluding the firm). $AnalystComove$ is the attention comovement measure computed by taking the logarithmic transformation of $Analyst\ R^2$ ($\ln(R^2/(1-R^2))$).$NewsComove$, $RetComove$ and $TOComove$ are computed using similar methods following Equations (4.3), (4.5) and (4.7). $Size$ is the market capitalization expressed in billions of dollars. $Price$ is the share price and BM is the book-to-market ratio measured at the end of June each year. $Turnover$ is average monthly share turnover. $NUMEST$ is the number of estimates underpinning the one-fiscal-year-ahead (FY1) earnings forecasts published in I/B/E/S. IO is the institutional ownership representing the percentage of a firm's share outstanding held by institutional investors. $SalesGrowth$ is the sales growth rate calculated as the current year sales divided by the prior year sales. ROA is the return on assets calculated each year. $StdROA$ is the standard deviation of ROA over the past 5 years. $ROAComove$ is the comovement in ROA estimated from regressing the firm's quarterly ROA on the equal-weighted style ROA over a 12-quarter window, and then take log transformation of the regression R^2. Panel B reports correlations among the key variables. Correlations are estimated in the cross section each year and then averaged over time. All variables are winsorized at the 1 and 99 percentiles each year before estimating summary statistics and correlations.

	Mean	Std. Dev.	Min	P25	P50	P75	Max
Panel A: Firm characteristics							
$Analyst$	6.0000	8.0000	1.0000	1.0000	3.0000	8.0000	97.0000
$News$	17.0000	28.1585	1.0000	3.0000	8.0000	19.0000	186.0000
$AnalystShock$	-0.1810	0.6722	-2.8811	-0.6748	-0.1178	0.3646	1.9095
$NewsShock$	-0.1766	0.7511	-4.4910	-0.6931	-0.0980	0.3673	3.4550
$Analyst\ R^2$	0.2222	0.2024	0.0000	0.0552	0.1732	0.3335	0.8595
$News\ R^2$	0.1301	0.1636	0.0000	0.0180	0.0713	0.1791	0.7936
$Turnover\ R^2$	0.1209	0.1358	0.0000	0.0150	0.0692	0.1864	0.5616
$Return\ R^2$	0.2048	0.1771	0.0000	0.0538	0.1617	0.3179	0.6870
$AnalystComove$	-1.9070	1.9972	-10.3500	-2.7617	-1.5093	-0.6447	4.4726
$NewsComove$	-2.9850	2.1777	-11.1622	-4.0056	-2.5758	-1.5357	2.1295
$TOComove$	-3.0515	2.2222	-11.3448	-4.1854	-2.5890	-1.4411	1.1552
$RetComove$	-1.9686	1.8386	-9.8688	-2.7747	-1.5759	-0.7140	1.4265
$Size$($billions)	3.6051	11.6549	0.0033	0.0936	0.4148	1.8774	154.0511
$Price$	24.4301	29.6857	-19.9250	5.4700	15.8000	33.1300	308.4000
BM	0.7250	0.8731	-2.4852	0.2927	0.5437	0.8917	12.0909
$Turnover$	0.1697	0.1807	0.0037	0.0527	0.1198	0.2191	1.9162
$NUMEST$	7.0000	7.0000	1.0000	2.0000	5.0000	10.0000	34.0000
IO	0.5332	0.3237	0.0001	0.2339	0.5768	0.8140	1.0000
$SalesGrowth$	1.1204	0.4379	0.0246	0.9635	1.0609	1.1810	5.9495
ROA	-0.0330	0.2322	-1.8475	-0.0284	0.0171	0.0671	0.3797
$StdROA$	0.0892	0.1588	0.0005	0.0136	0.0369	0.0954	1.3285
$ROAComove$	-2.2908	1.5520	-6.3903	-3.2473	-2.1785	-1.2254	0.8965

Table 4.1 Continued

		1	2	3	4	5	6	7	8	9	10	11	12	13
Panel B: Correlation matrix														
1	$AnalystComove$	1												
2	$NewsComove$	0.2069	1											
3	$TOComove$	0.0558	0.0343	1										
4	$RetComove$	0.0137	0.0004	0.3246	1									
5	$ROAcomove$	0.0057	0.0113	0.0275	0.0438	1								
6	$Size$	-0.0242	-0.0417	0.4503	0.5167	0.0096	1							
7	BM	-0.0236	0.0493	-0.1171	-0.1242	0.0265	-0.2979	1						
8	$Turnover$	-0.0633	-0.1032	0.0693	0.0949	0.0259	0.1804	-0.1656	1					
9	$NUMEST$	-0.0222	-0.0373	0.3485	0.3484	0.0316	0.7807	-0.2095	0.2661	1				
10	IO	0.0047	-0.0293	0.2608	0.3892	0.0155	0.5507	-0.1399	0.1354	0.4484	1			
11	$SalesGrowth$	-0.0020	-0.0183	-0.0120	-0.0053	-0.0064	-0.0116	-0.0406	0.0294	-0.0101	-0.0248	1		
12	ROA	-0.0256	0.0465	0.0852	0.1083	0.0027	0.2054	0.0565	-0.0649	0.1687	0.2129	-0.1037	1	
13	$StdROA$	-0.0001	-0.0247	-0.0464	-0.0419	0.0032	-0.0763	-0.0966	0.0591	-0.0698	-0.0956	0.0400	-0.1151	1

4.5 Empirical results

4.5.1 Style performance and investor attention

Barberis and Shleifer's (2003) style investing model is built on the assumption of investors' relative style chasing behaviour. We begin our empirical analysis with an examination of whether prior style performance affects investor attention. Hypothesis 1 predicts that, since investors allocate funds based on styles' past performance and shift funds from poorly performed styles to well performed styles, style portfolios with extreme past returns are expected to attract more investor attention.

To test the hypothesis, we regress style-level attention shock on lagged absolute style returns while controlling for style average size and book-to-market as specified in Equation (4.8). The results are reported in Table 4.2. In regression (1), $AnalystShock$ is used as the proxy for attention shock. It shows that style-level $AnalystShock$ is significantly and positively related to absolute style returns over the past 1 month ($|Ret_{S,t-1}|$), 2 to 3 months ($|Ret_{S,t-3,t-2}|$), and 4 to 6 months ($|Ret_{s,t-6,t-4}|$). The coefficient on $|Ret_{S,t-1}|$ is equal to 0.4262, suggesting a 1% increase in the absolute return over the past one month translates into a 0.43% increase in attention shock. The magnitude of the coefficient drops to 0.3184 on $|Ret_{S,t-3,t-2}|$ and 0.1931 on $|Ret_{s,t-6,t-4}|$). Similar results are documented in regression (2), when news coverage shock is used to proxy for attention shock. Since style returns are expressed in absolute values in our analysis, the positive relation between attention shock and absolute style returns suggests that styles with more extreme performances attract more investor attention. This finding is consistent with the existing literature which finds that investors are likely to perceive extreme styles as special portfolios (e.g., Kumar 2009). Also, it provides supportive evidence for Barberis and Shleifer's (2003) style chasing model.

Prior studies suggest that firm-specific attention is positively related to the firm's abnormal stock returns (e.g., Ben-Rephael et al., 2017; Drake et al., 2015). Thus, a question of interest is whether the observed performance-driven attention at the style level is distinct from the firm-level attention documented in previous studies. To address this question, we regress firm-level attention on prior style returns while controlling for individual stock returns as specified in Equation (4.9). Table 4.3 shows that $|Ret_{s,t-3,t-2}|$, $|Ret_{s,t-6,t-4}|$, and $|Ret_{s,t-12,t-7}|$ remain significantly positive for both measures of attention shock. This suggests that style returns have predictive power for firm-level attention even after controlling for individual stock returns. By contrast, the relation between firm-level attention and individual stock returns is weak, varying with time-horizon and attention measures. This further supports the notion that investors invest at the style level rather than at the individual stock level.

In summary, consistent with hypothesis 1, Tables 4.2 and 4.3 show that prior style performance plays an important role in driving both style-level and firm-level attention. The results provide empirical evidence for Barberis and Shleifer's (2003) style investing and style performance chasing prediction.

Table 4.2 Style-level attention and prior style returns

This table presents the impact of prior style returns on style-level investor attention. Style-level attention shock is regressed on lagged absolute style returns, style average market capitalization ($Size_s$) and book-to-market ratios(BM_s). Attention shock is measured by the detrended analyst forecast revisions and detrended news coverage as described in Equations (4.1a) and (4.1b). Style-level attention shock is calculated by averaging the individual attention shock across stocks within the same style. Ret_s, $Size_s$, and BM_s are the equal-weighted average of stock returns, market capitalization, and book-to-market of individual stocks within the style. $Size$ and BM are natural logged. The table reports the regression results across style-months with style and month fixed effects. T-statistics based on style-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $AnalystShock_{s,t}$	(2) $NewsShock_{s,t}$
$Size_{s.t}$	0.1467***	0.0988**
	(3.40)	(2.65)
$BM_{s,t}$	0.0134	0.0268
	(0.62)	(1.09)
$\|Ret_{s,t-1}\|$	0.4262**	0.4512**
	(2.52)	(2.19)
$\|Ret_{s,t-3,t-2}\|$	0.3184***	0.2359***
	(3.04)	(2.83)
$\|Ret_{s,t-6,t-4}\|$	0.1931**	0.0466
	(2.76)	(0.74)
$\|Ret_{s,t-12,t-7}\|$	-0.0268	-0.1339*
	(-0.67)	(-1.88)
Constant	-1.4510***	-1.0737**
	(-2.89)	(-2.70)
Style fixed-effects	yes	yes
Month fixed-effects	yes	yes
Observations	3,799	3,775
R-squared	0.7744	0.5825

Table 4.3 Firm-level attention and prior style returns

This table presents the impact of prior style returns on firm-level investor attention. As specified in Equation (4.7), firm-level attention shock is regressed on lagged absolute style returns ($|Ret_S|$), lagged absolute stock returns ($|Ret_i|$), firm size ($Size_i$) and firm book-to-market ratio (BM_i). Attention shock is measured by the detrended analyst forecast revisions and detrended news coverage as described in Equations (4.1a) and (4.1b). Style return is the equal-weighted stock returns within the style. $Size$ and BM are natural logged. The table reports the regression results across firm-months with style and month fixed effects. T-statistics based on firm-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1)	(2)
	$AnalystShock_{i,t}$	$NewsShock_{i,t}$
$Size_{i.t}$	0.0065**	-0.0040***
	(2.54)	(-2.82)
$BM_{i,t}$	-0.0399***	-0.0259***
	(-13.94)	(-11.13)
$\lvert Ret_{S,t-1}\rvert$	0.0516	0.1729**
	(0.72)	(2.57)
$\lvert Ret_{S,t-3,t-2}\rvert$	0.0864**	0.1277***
	(2.05)	(2.98)
$\lvert Ret_{S,t-6,t-4}\rvert$	0.1872***	0.1872***
	(6.12)	(5.92)
$\lvert Ret_{S,t-12,t-7}\rvert$	0.0486**	0.1361***
	(2.48)	(6.58)
$\lvert Ret_{i,t-1}\rvert$	-0.2004***	0.0091
	(-11.41)	(0.79)
$\lvert Ret_{i,t-3,t-2}\rvert$	0.0147*	0.0496***
	(1.83)	(6.90)
$\lvert Ret_{i,t-6,t-4}\rvert$	0.0004	-0.0142**
	(0.06)	(-2.30)
$\lvert Ret_{i,t-12,t-7}\rvert$	0.0222***	-0.0087
	(6.62)	(-1.36)
Constant	0.0205	0.6166**
	(0.96)	(2.37)
Style fixed-effects	Yes	Yes
Month fixed-effects	Yes	Yes
Observations	328,564	430,664
R-squared	0.1313	0.0576

4.5.2 Investor attention and asset comovement

4.5.2.1 Style attention and attention comovement

One of the important asset-pricing implications of Barberis and Shleifer (2003) is that style investing generates within-style return comovement. In their framework, within-style return comovement arises from investors' concentrated demand for a certain group of securities.[58] It is plausible that, when investors systematically seek information about a certain group of stocks, their attention comoves. This leads to common movement in trading activity and stock prices. As such, we expect that attention-grabbing styles are associated with strong attention comovement as well as strong comovement in trading activity and stocks returns.

We use portfolio sorts as a preliminary test for this conjecture. At the end of June of each year t, we sort styles into quintile portfolios based on their average attention shock, measured by $AnalystShock$ or $NewsShock$, over the period from July of year t-1 to June of year t.[59] We then estimate the attention comovement, turnover comovement and return comovement in each attention quintile over the same period and report the results in Table 4.4.

Panel A presents the comovement in different attention groups formed on $AnalystShock$. Consistent with hypothesis 2, investor attention comoves more among high attention-grabbing styles. $AttnComove$ in the top attention quintile is 0.25 higher than the $AttnComove$ in the bottom attention quintile. Also, more attention-grabbing styles exhibit stronger return comovement than less attention-grabbing styles. $RetComove$ is equal to -0.8634 in the top attention quintile compared to -1.4087 in the bottom attention quintile, and the difference is statistically significant at the 1% level. Similarly, turnover comovement is significantly higher in the high attention quintile than in the low attention quintile. However, we did not observe a significant difference in either attention comovement or return comovement between the high and low attention quintiles in Panel B, when $NewsShock$ is used to form attention groups. Since $NewsShock$ is more likely to capture the attention shocks of retail investors, the results suggest that the style level return comovement is primarily driven by the style-induced trading of institutional investors. This finding is consistent with the fact that institutional investors are the main style investors.

[58] In line with this view, Kumar (2009) finds that the strength of return comovement within a particular style increases when investors move into or out of the style with greater intensity.
[59] We choose the 12-month window from July of year *t-1* to June of *t* to match up with the style formation period.

Table 4.4 Attention comovement, turnover comovement and return comovement of portfolios sorted on style attention

This table presents attention comovement, turnover comovement and return comovement in style portfolios sorted on attention shock. At the end of June of each year t, we sort styles into quintiles based on their average attention shock, measured by $AnalystShock$ or $NewsShock$, over the period from July of year t-1 to June of year t. We report the attention comovement ($AttnComov$), turnover comovement ($TOComove$) and return comovement ($RetComove$) for each attention group, respectively. Attention comovement of an individual stock captures the degree to which firm-specific attention is explained by the attention paid to its style, estimated using the R^2 from Equation (4.2). For each attention quintile, we first compute the equal-weighted attention comovement of each style, and then take the average of each style's attention comovement in that quintile. Turnover comovement and return comovement are computed in the same way. Panel A presents the average attention comovement, turnover comovement and return comovement in $AnalystShock$ quintiles. Panel B presents the results in $NewsShock$ quintiles. The difference between the high and low attention quintiles is also reported, along with t-statistics in parentheses.

	$AttnComove$	$TOComove$	$RetComove$
Panel A: Attention comovement, turnover comovement and return comovement in $AnalystShock$ groups			
Low attention	-3.1735	-2.6013	-1.4087
2	-3.1378	-2.7465	-1.5460
3	-3.1285	-2.6977	-1.5245
4	-3.1910	-2.2099	-1.2232
High attention	-2.9231	-1.5533	-0.8634
High - Low	0.2504***	1.0480***	0.5453***
t-Statistic	(3.41)	(8.22)	(3.97)
Panel B: Attention comovement, turnover comovement and return comovement in $NewsShock$ groups			
Low attention	-2.5460	-2.0564	-1.0975
2	-2.4432	-2.4651	-1.3786
3	-2.4737	-2.4015	-1.3741
4	-2.4403	-2.5701	-1.4978
High attention	-2.4261	-2.3156	-1.2177
High - Low	0.1199	-0.2592	-0.1202
t-Statistic	(1.05)	(-1.56)	(-0.78)

4.5.2.2 Attention comovement and return comovement

The previous section provides style-level evidence for the association between investor attention and asset comovement. In this section, we investigate the impact of attention comovement on return comovement and turnover comovement at the individual stock level. Prior studies suggest that attention-constrained investors tend to allocate more attention to broader level information than to firm-specific information, and excessive comovement is driven by investors' inattention to firm-specific information (e.g., Peng and Xiong, 2006). When firm-specific attention to a stock is, to a large extent, explained by general attention paid to its style, its returns should comove with the style returns. Building on this view, we conjecture that a stock's attention comovement with the style is positively associated with the return comovement as well as the turnover comovement.

We test this conjecture by performing different specifications of Equation (4.10), and present the results in Table 4.5. Panel A reports the coefficients from regressing return comovement on the contemporaneous attention comovement. Consistent with the conjecture, we document a significant positive relation between attention comovement and return comovement. Coefficients on control variables suggest that return comovement is higher among large size, high book-to-market and high institutional ownership stocks. Unsurprisingly, stocks whose fundamentals comove more with other stocks within the same style exhibit stronger return comovement. Coefficients on ROA comovement ($ROACom$) are significantly positive at the 1% level in all the regressions. More importantly, the finding on a positive relation between attention comovement and return comovement after controlling for fundamental comovement suggests that, information flows plays an incremental role in explaining return comovement that is beyond fundamentals. Panel B presents the results from the regression of turnover comovement on attention comovement. The results support a positive relation between attention comovement and turnover comovement.

Previous studies show that extreme returns and trading activity can shift investor attention to a particular stock. Thus, the relation between attention and market activity may be endogenous. To address this concern, we test whether attention comovement has predictive power on return comovement and turnover comovement. To do so, we modify the specification in Equation (4.10) to examine the lead-lag relation between return (or turnover) comovement and attention comovement, also controlling for the lagged return comovement. The results are presented in Table 4.6.

Panel A suggests that a stock's attention comovement predicts its return comovement over the subsequent period. Coefficients on lagged attention comovement are significantly positive in all regressions. In addition, within-style return comovement demonstrates a low level of persistence over time, as coefficients on lagged return comovement are below 0.30 in all regressions. Similarly, Panel B reveals a significant positive relation between attention comovement and the subsequent turnover comovement. To conserve space and focus on the marginal effects of attention comovement, the table supresses coefficients of the controls. In the untabulated results, coefficients on the control variables remain similar to those in Table 4.5.

Overall, in line with Hypothesis 3a, results in this section indicate that attention comovement is positively related to return (turnover) comovement after controlling for comovement in fundamentals. Thus, our finding provides an information flow explanation for the excess comovement among stocks in the same style.

Table 4.5 Return comovement , turnover comovement and attention comovement

This table presents how attention comovement is associated with return comovement and turnover comovement. Panel A reports the contemporaneous relation between attention comovement and return comovement and Panel B reports the relation between attention comovement and turnover comovement. $RetComove_{i,t}$ represents the return comovement for stock *i* in year *t*, measured using the R^2 from regressing weekly returns of stock *i* on equal-weighted style returns and taking the log transformation. $TOComove_{i,t}$ represents the turnover comovement for stock *i* in year *t*, measured using the R^2 from regressing weekly turnover of stock *i* on equal-weighted style turnover and taking the log transformation. $AnalystComove$ and $NewsComove$ are measures of attention comovement based on the number of analyst forecast revisions and news coverage, respectively (as described in Equation (4.3)). $ROACom$ is the ROA comovement proxying for the firm's fundamental comovement with its style. Other control variables include market capitalization ($Size$), book-to-market ratio (BM), share price($Price$), absolute stock return ($AbsRet$), stock turnover ($Turnover$), analyst coverage($Analyst$), the fraction of shares outstanding held by institutional investors(IO), return on assets (ROA), and standard deviation of return on assets ($StdROA$). Size, book-to-market, price, analysts and institutional ownership are natural logged. The table reports the panel regression results with style and year fixed effects. T-statistics based on firm-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	Panel A: $RetComove_{i,t}$		Panel B: $TOComove_{i,t}$	
	(1)	(2)	(3)	(4)
$AnalystComove_{i,t}$	0.0197***		0.0820***	
	(5.56)		(15.84)	
$NewsComove_{i,t}$		0.0104***		0.0721***
		(3.39)		(14.81)
$ROA_{i,t}$	-0.1501***	-0.2181***	0.0274	-0.0878
	(-2.77)	(-4.14)	(0.33)	(-1.09)
$Size_{i,t}$	0.2984***	0.3792***	0.2636***	0.3813***
	(18.36)	(24.24)	(13.21)	(18.96)
$BM_{i,t}$	0.0659***	0.0850***	-0.0014	0.0188
	(4.95)	(6.44)	(-0.06)	(0.85)
$SalesGrowth_{i,t}$	-0.0488***	-0.0697***	-0.0576*	-0.0437
	(-2.83)	(-3.86)	(-1.89)	(-1.41)
$IO_{i,t}$	0.0664***	0.0763***	0.0367***	0.0360***
	(6.87)	(8.57)	(2.76)	(2.90)
$Analyst_{i,t}$	-0.0455**	-0.0044	0.0393	0.0640**
	(-2.33)	(-0.23)	(1.50)	(2.53)
$StdROA_{i,t}$	-0.0826	-0.1437**	-0.0010	-0.0411
	(-1.22)	(-2.23)	(-0.01)	(-0.37)
$AbsRet_{i,t}$	-0.2285***	-0.2388***	0.2504***	0.3130***
	(-11.72)	(-12.71)	(7.52)	(9.97)
$Turnover_{i,t}$	-0.4576***	-0.1661***	-1.0551***	-0.7556***
	(-7.35)	(-2.76)	(-12.11)	(-8.79)
$Price_{i,t}$	-0.0069	-0.0348**	-0.0159	-0.0477***
	(-0.49)	(-2.54)	(-0.88)	(-2.60)
$ROACom_{i,t}$	0.0351***	0.0326***	0.0440***	0.0349***
	(8.60)	(7.90)	(6.82)	(5.25)
Constant	-3.1291***	-3.5082***	-4.7528***	-5.2565***
	(-38.22)	(-44.40)	(-40.89)	(-45.69)
Style fixed-effects	Yes	Yes	Yes	Yes
Year fixed-effects	Yes	Yes	Yes	Yes
Observations	34,994	33,848	36,649	35,984
R-squared	0.2815	0.320	0.2633	0.2647

Table 4.6 Return comovement, turnover comovement and attention comovement - Lead-lag relations

This table presents how attention comovement is associated with future return comovement and turnover comovmenet. Panel A reports the lead-lag relation between attention comovement and return comovement and Panel B reports the lead-lag relation between attention comovement and turnover comovement. $RetComove_{i,t}$ represents the return comovement for stock *i* in year *t* and $TOComove_{i,t}$ represents the turnover comovement for stock *i* in year *t*. $AnalystComove_{i,t-1}$ and $NewsComove_{i,t-1}$ are measures of attention comovement for stock *i* in year *t-1* based on the number of analyst forecast revisions and news coverage, respectively (as described in Equation (4.3)). $ROACom$ is the ROA comovement proxying for the firm's fundamental comovement with its style. $Controls$ include other control variables specified in Table 4.5, including market capitalization ($Size$), book-to-market ratio (BM), share price($Price$), absolute stock return ($AbsRet$), stock turnover ($Turnover$), analyst coverage($Analyst$), the fraction of shares outstanding held by institutional investors(IO), return on assets (ROA), and standard deviation of return on assets ($StdROA$). The table reports the panel regression results with style and year fixed effects. T-statistics based on firm-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	Panel A: $RetComove_{i,t}$		Panel B: $TOComove_{i,t}$	
	(1)	(2)	(3)	(4)
$AnalystComove_{i,t-1}$	0.0090***		0.0635***	
	(2.83)		(12.10)	
$NewsComove_{i,t-1}$		0.0078***		0.0305***
		(2.74)		(6.36)
$RetComove_{i,t-1}$	0.2967***	0.2877***		
	(42.05)	(39.34)		
$ToComove_{i,t-1}$			0.0854***	0.0805***
			(13.31)	(12.66)
$ROACom_{i,t-1}$	0.0224***	0.0216***	0.0353***	0.0340***
	(5.83)	(5.47)	(5.41)	(5.00)
$Controls_{i,t-1}$	Yes	Yes	Yes	Yes
Style fixed-effects	Yes	Yes	Yes	Yes
Year fixed-effects	Yes	Yes	Yes	Yes
Observations	32,589	31,278	35,720	35,116
R-squared	0.3316	0.355	0.2769	0.2607

4.5.2.3 Changes in comovement after style reclassification

The positive relation between within-style attention comovement and return comovement provides empirical support for the behavioural explanations of excess price comovement. In particular, investors view firms in the context of style categories, generating excess comovement in assets within the same style. To support this finding, we further test how a stock's comovement in attention, return and turnover changes following its style reclassification. If price comovement arises from category investment, we expect that style reclassification leads to changes in asset comovement.

We study a subset of stocks experiencing year-on-year style changes over the 2003–2016 period. There are 25,815 out of 53,685 firm-years in our sample that experience style changes. For each event stock, we estimate its slope coefficients on the old style and new style over the 11-month pre-event period and 11-month post-event period, respectively, as described in Equation (4.11). Table 4.7 Panel A presents the mean coefficients on the pre-reclassification style ($\overline{\beta_{OldStyle}}$) and post-reclassification style ($\overline{\beta_{NewStyle}}$) over the pre-event and post-event windows, separately, as well as the changes in the slope coefficients.

We first report the findings on changes in attention comovement after style reclassification using $Analyst$ as the attention measure. We show that style reclassification is associated with a significant increase in the coefficient on the new style and a significant decrease in the coefficient on the old style. On average, the coefficient on the new-style attention ($\overline{\beta_{NewStyle}^{Analyst}}$) increases by 0.0633 and the coefficient on the old-style attention ($\overline{\beta_{OldStyle}^{Analyst}}$) drops by -0.0330. The results are similar when $News$ is used as the attention measure.

We then report changes in return comovement and turnover comovement after style reclassification. After entering a new style, a stock's return comovement with that style increases significantly. The average coefficient on news-style return ($\overline{\beta_{NewStyle}^{Ret}}$) equals 0.0872 for the pre-event window, and it increases to 0.1066 for the post-event window. The difference of 0.0194 is statistically significant at the 5% level. Also, the stock's comovement in turnover with the old (new) style falls (rises), with the coefficient on the old-style drops by -0.0949 and on the new-style rises by 0.0863.

In Panel B, we focus on stocks experiencing more significant style changes and request a minimum of two-style migration.[60] Consistent with Panel A, a stock's comovement in attention, returns and turnover with the old (new) style decreases (increases) following the style reclassification. However, the magnitude of slope changes is much larger compared to those in Panel A. For example, the coefficient on new-style turnover ($\overline{\beta_{NewStyle}^{Turnover}}$) increases by 0.1466 and on old-style turnover ($\overline{\beta_{OldStyle}^{Turnover}}$) decreases by -0.1906 , compared to 0.0863 and -0.0949 documented in Panel A.

Overall, the observed changes in slope coefficients following style reclassification are consistent with Hypotheses 3b and 3c. When a stock is classified into a certain style, it enters a category used by many investors, and is influenced by fund flows in and out of that category. These fund flows raise the correlation of the included stock's return with the returns of other stocks within the same style. Our results provide further evidence for the category and habitat views on return comovement.

[60] For example, a stock moving from style 1×1 (bottom size quintile and bottom book-to-market quintile) to style 1×3 (bottom size quintile and the 3rd book-to-market quintile) is considered as a two-style migration. A stock moving from style 1×1 to style 2×2 (the 2nd size quintile and the 2nd book-to-market quintile) is also considered as a two-style migration.

Table 4.7 Changes in attention comovement, return comovement, and turnover comovement following style reclassifications

This table presents changes in attention, return and turnover comovement following a stock's style change. At the end of June each year over the sample period, we form 5×5 size and book-to-market style portfolios. The sample includes stocks experiencing year-to-year style changes. For each event stock *i*, the bivariate model

$$Attn_{i,t} = \alpha_i + \beta_{i,OldStyle}Attn_{OldStyle,t} + \beta_{i,NewStyle}Attn_{NewStyle,t} + \varepsilon_{i,t}$$

are separately estimated for the pre- and post-event period. $Attn$ is the investor attention measured by the number of analyst forecast revisions ($Analyst$) and news coverage ($News$). We examine the mean changes in the slopes, $\overline{\Delta\beta_{OldStyle}}$ and $\overline{\Delta\beta_{NewStyle}}$. The pre- and post-event estimation window are [-11,-1] and [+1,+11] months. We use the similar bivariate model to estimate changes in return comovement and turnover comovement. Panel A reports the slopes from the pre-event period and post-event period for the entire sample. Panel B reports the results for stocks with a minimum of two-quintile migration. Changes in the slopes are reported along with t-statistics in parentheses. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	$\overline{\beta^{Analyst}_{OldStyle}}$	$\overline{\beta^{Analyst}_{NewStyle}}$	$\overline{\beta^{News}_{OldStyle}}$	$\overline{\beta^{News}_{NewStyle}}$	$\overline{\beta^{Ret}_{OldStyle}}$	$\overline{\beta^{Ret}_{NewStyle}}$	$\overline{\beta^{Turnover}_{OldStyle}}$	$\overline{\beta^{Turnover}_{NewStyle}}$
Panel A: Changes in comovement around style reclassification								
Pre-event	0.4380	0.3825	0.4181	0.3602	0.0900	0.0872	0.4550	0.4147
Post-event	0.4049	0.4458	0.3742	0.4594	0.0812	0.1066	0.3601	0.5011
$\overline{\Delta\beta}$	-0.0330*	0.0633***	-0.0438**	0.0992***	-0.0089	0.0194**	-0.0949***	0.0863***
	(-1.77)	(3.24)	(-2.19)	(4.85)	(-1.18)	(2.43)	(-4.48)	(3.83)
Panel B: Changes in comovement around large style migration								
Pre-event	0.4169	0.3473	0.4160	0.3526	0.1154	0.0842	0.5805	0.4912
Post-event	0.3528	0.4454	0.3324	0.4941	0.0668	0.1284	0.3898	0.6377
$\overline{\Delta\beta}$	-0.0641**	0.0981***	-0.0837***	0.1414***	-0.0486***	0.0442***	-0.1906***	0.1466***
	(-2.32)	(3.20)	(-2.78)	(4.47)	(-3.21)	(2.86)	(-5.46)	(3.81)

4.5.3 Style attention and the cross-section of returns

4.5.3.1 Style attention and style returns

After the establishment of within-style attention comovement and its relation with return comovement, this section takes a further step to investigate whether style-level attention contributes to the variation in the cross-section of style returns. Hypothesis 4a predicts that high style-level attention is associated with a short-term increase and long-term reversal in style returns.

At the end of each month, we sort style portfolios into quintiles based on their level of attention shock, and calculate returns for each attention quintile. We consider different portfolio formation and holding periods to capture the dynamic relation between style attention and style returns. First, we sort styles on attention shock in each month and examine the contemporaneous style returns.[61] We then consider strategies that sort styles based on attention shock of prior one month and hold the portfolio for one month (1 month – 1 month strategy), as well as strategies that use holding periods and formation periods of 3 months (3 month – 3 month strategy) , 6 months (6 month – 6 month strategy) and 12 months (12 month – 12 month strategy). If high attention leads to a temporary price increase and a subsequent price reversal, high-attention styles should outperform low-attention styles in short-horizon strategies, and underperform low-attention styles in long-horizon strategies.

For each strategy, we sort styles into quintiles based on style-level attention shock over the formation period and then examine the returns of the portfolio over the holding period. Following the standard practice in the literature, we skip a month between the formation period and the subsequent holding period to ensure that microstructure biases do not affect our tests. For example, for the 3 month – 3 month strategy at month t, portfolio 1 (5) consist of the quintile of styles experiencing the least (most) attention shock over months t-3 to t-1. For each quintile, we first compute the equal-weighted return of each style in the portfolio, and then take an equal-weighted average of each style's return in that portfolio from months t+1 to t+3.[62] Table 4.8 displays the average returns for each quintile over different combinations.

[61] Similar to the earlier analysis, attention shock of an individual stock is measured by the detrended analyst forecast revisions and detrended news coverage as described in Equations (4.1a) and (4.1b). Style-level attention shock is calculated by averaging the individual attention shock across stocks within the same style.

[62] In the untabulated analysis, we also compute the value-weighted style returns, and then equally weight each style's return in the attention quintile. Results are qualitatively similar.

Panel A presents the results using the number of analyst revisions as the attention measure ($AnalystShock$). Style attention is positively associated with the contemporary style returns. On average, styles attracting the most attention in the month outperform styles attracting the least attention by 46bps, and the difference is statistically significant at the 5% level. When both the estimation and holding periods are extended, there is evidence of return reversals. Specifically, high attention-grabbing styles over the past one month underperform the low attention-grabbing styles by 44bps over the subsequent one month. High-attention styles over the past three months underperform the low-attention styles by 18bps over the following three months. Results are qualitatively similar for the 6 month – 6 month and 12 month – 12 month strategies. The return differences between the low and high attention quintiles for the two strategies are 13 and 16 bps, respectively, and are both statistically significant.

Similar patterns are documented in Panel B when news coverage is used as the proxy for attention. On average, high $NewsShock$ styles outperform low $NewsShock$ styles by 36 bps in the concurrent month. Over the long run, high $NewsShock$ styles are associated with lower returns. For example, the high-minus-low return spread is 14 bps for the 6 month – 6 month strategy, and 10 bps for the 12 month - 12 month strategy. Overall, results in Table 4.8 support hypothesis 4a that high style-level attention is associated with high style returns in the short run and reversal in the long run.

Table 4.8 Returns on portfolios sorted on style-level attention shocks

This table presents style returns on portfolios sorted on investor attention. At the end of each month over the sample period from January 2003 to March 2016, styles are sorted into quintiles based on their attention shock in the current month, and over the past 1, 3, 6, and 12 months. The styles experiencing the least (most) attention shocks are placed in portfolio 1 (5). Returns to the quintile portfolios are equally-weighted. The average monthly returns of these portfolios over different holding period are presented in this table. The difference in returns between quintile 5 and 1 is reported (High - Low), along with t-statistics in parentheses. Panel A and B report the results using $AnalystShock$ and $NewsShock$ as the proxy for attention shock, respectively. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	Low attention	2	3	4	High attention	High - Low	t-Statistic
Panel A: Style returns on portfolios sorted on $AnalystShock$							
Current month	0.0070	0.0098	0.0095	0.0104	0.0115	0.0046**	(2.58)
1 month - 1 month	0.0133	0.0120	0.0111	0.0108	0.0089	-0.0044***	(-2.98)
3 month - 3 month	0.0102	0.0094	0.0081	0.0083	0.0084	-0.0018*	(-1.89)
6 month - 6 month	0.0083	0.0074	0.0064	0.0072	0.0070	-0.0013**	(-2.13)
12 month - 12 month	0.0076	0.0068	0.0072	0.0066	0.0060	-0.0016***	(-4.12)
Panel B: Style returns on portfolios sorted on $NewsShock$							
Current month	0.0078	0.0088	0.0108	0.0104	0.0113	0.0036***	(2.71)
1 month - 1 month	0.0111	0.0113	0.0095	0.0094	0.0092	-0.0019	(-1.18)
3 month - 3 month	0.0079	0.0085	0.0090	0.0088	0.0072	-0.0007	(-0.75)
6 month - 6 month	0.0075	0.0075	0.0078	0.0075	0.0060	-0.0014**	(-2.30)
12 month - 12 month	0.0067	0.0074	0.0074	0.0072	0.0057	-0.0010***	(-2.71)

4.5.3.2 Style attention and autocorrelations in style returns

Barberis and Shleifer (2003) predict that style investing generates autocorrelations in style returns. In particular, investors constantly buy outperforming styles and sell underperforming style. The concentrated demand pushes prices beyond fundamentals, creating short-term momentum and long-term reversals in style prices. If investor attention is an important source of style investing, it should contribute to the consequent return patterns. Accordingly, we investigate how style-level attention is related to autocorrelations in style returns.

To test this, we examine the relation between style attention and future style returns conditional on prior style performance. We argue that investors allocate funds to attention-grabbing outperforming styles and withdraw funds from attention-grabbing underperforming styles, which temporarily pushes up the prices of outperforming styles and pushes down the prices of underperforming styles from fundamentals. As a consequence, high style-level attention should positively forecast style returns over short run and negatively forecast returns over long run for the outperforming styles, and signs should be opposite for the underperforming styles.

We use sequential portfolio sorts to conduct the investigation. At the end of each month, styles are sorted into quintiles on returns over the past one month, three months, six months or 12 months. We refer to styles in the bottom (top) return quintile as loser (winner). Within each winner and loser group, styles are further sorted into quintiles based on their attention shocks over the same formation period. We calculate the average monthly returns over the subsequent period for each attention quintile. Results are presented in Table 4.9.

Panel A uses $AnalystShock$ as the attention measure. We begin by examining the 1 month – 1 month strategy. For the underperforming styles (loser), top attention-grabbing styles over the past one month underperform bottom attention-grabbing styles by 54 bps in the following month. By contrast, among the outperforming styles (winner), the top attention quintile outperforms the bottom attention quintile by 58 bps in the subsequent month. This finding is consistent with the explanation that concentrated demand by style switchers pushes the prices of outperforming styles up and pushes the prices of underperforming styles down. Investigation of the longer-horizon attention-return relation demonstrates a strong price reversal among outperforming styles. For the winner group, the high attention quintile underperforms the low attention quintile by 25 bps, 19 bps and 16 bps for 3 month – 3 month strategy, 6 month – 6 month strategy, and 12 month – 12 month strategy. The reversal pattern in the loser group is not as strong as in the winner group. Nevertheless, the economic significance of the return difference between the top and bottom attention quintiles in the loser group decreases with the passage of time. The return difference decreases from 26 bps for 3 month – 3 month strategy to 20 bps for 6 month – 6 month strategy, and further to 16 bps for 12 month – 12 month strategy, indicating a reverse pattern in returns.

Panel B presents the results using $NewsShock$ as the attention measure. The findings are consistent with those in Panel A. For the outperforming group, high attention-grabbing styles outperform low attention-grabbing styles over the short run, and underperform low attention styles over the long run. For example, high attention styles outperform low attention styles by 25 bps for 3 month – 3 month strategy, but underperform low attention styles by 18 bps for 6 month – 6 month strategy. For the underperforming groups, high attention styles underperform low attention styles by 54 bps for 1 month – 1 month strategy. The return difference reduces significantly to 16 bps for 12 month – 12 month strategy.

Overall, Table 4.9 shows that style-level attention contributes to autocorrelation in style returns. Investors switch capital from attention-grabbing underperforming styles to attention-grabbing outperforming styles, temporarily pushing prices away from fundamentals.

Table 4.9 Returns on portfolios sorted on style-level attention shocks in winner and loser groups

This table examines style returns on portfolios sorted on style-level attention shock within the winner and loser groups, separately. At the end of each month, styles are sorted into quintiles based on their returns over the past 1, 3, 6, and 12 months. For each style, we compute the equal-weighted returns. Styles with the lowest (highest) returns are placed in loser (winner) group. Within the loser (or winner) group, styles are further sorted into quintiles based on the level of attention shocks over the same formation period. The average monthly returns of these portfolios over different holding periods are presented in this table. Within the winner (or loser) group, the return difference between high- and low-attention quintiles is reported (High - Low), along with t-statistics in parentheses. Panel A reports the results using $AnalystShock$ as the proxy for attention shock, and panel B reports the results using $NewsShock$ to proxy for attention shock. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

		Low	2	3	4	High	High - Low	t-Statistic
Panel A: Returns on portfolios sorted on style $AnalystShock$ in winner and loser groups								
1 month – 1 month	Loser	0.0102	0.0092	0.0074	0.0076	0.0048	-0.0054**	(-2.21)
	Winner	0.0108	0.0093	0.0141	0.0123	0.0166	0.0058*	(1.95)
3 month – 3 month	Loser	0.0089	0.0085	0.0074	0.0058	0.0063	-0.0026*	(-1.75)
	Winner	0.0095	0.0095	0.0096	0.0084	0.0070	-0.0025*	(-1.84)
6 month – 6 month	Loser	0.0083	0.0075	0.0059	0.0074	0.0063	-0.0020**	(-2.15)
	Winner	0.0075	0.0078	0.0080	0.0064	0.0056	-0.0019*	(-1.93)
12 month – 12 month	Loser	0.0067	0.0056	0.0069	0.0070	0.0051	-0.0016**	(-1.98)
	Winner	0.0068	0.0059	0.0069	0.0060	0.0052	-0.0016**	(-2.46)
Panel B: Returns on portfolios sorted on style $NewsShock$ in winner and loser groups								
1 month – 1 month	Loser	0.0122	0.0073	0.0068	0.0085	0.0064	-0.0058**	(-2.26)
	Winner	0.0103	0.0116	0.0135	0.0111	0.0125	0.0023	(1.07)
3 month – 3 month	Loser	0.0061	0.0077	0.0080	0.0078	0.0064	0.0002	(0.17)
	Winner	0.0083	0.0117	0.0102	0.0094	0.0108	0.0025*	(1.68)
6 month – 6 month	Loser	0.0089	0.0081	0.0065	0.0069	0.0062	-0.0027***	(-2.85)
	Winner	0.0079	0.0072	0.0062	0.0073	0.0061	-0.0018*	(-1.89)
12 month – 12 month	Loser	0.0073	0.0056	0.0061	0.0064	0.0057	-0.0016*	(-1.84)
	Winner	0.0064	0.0057	0.0066	0.0059	0.0062	-0.0002	(-0.20)

4.5.3.3 Style attention and individual stock returns

Prior studies show that firm-level attention predicts higher stock prices in the short run and price reversal in the long run (e.g., Barber and Odean, 2008; Da et al., 2011). A natural question is whether style-level attention also influences firm-level returns after controlling for firm-level attention. If investors categorize stocks into styles when making investment decisions, we expect the impact of style-level attention on stock returns remains after controlling for firm-specific attention. To explore this conjecture, we first sort stocks into quintiles based on firm-level attention. Within each firm-level attention quintile, we further sort stocks into quintiles based on their style-level attention, resulting in 25 portfolios. As in Table 4.8, we start with examining the contemporaneous relation between style attention and stock returns, and then consider four formation periods (1, 3, 6, and 12 month) and four holding periods (1, 3, 6, and 12 month). Within each portfolio, we calculate the equal-weighted returns in each month, and report the average monthly returns over the holding period.[63]

Panel A presents the results using $AnalystShock$ as the attention measure. Results on the contemporaneous relation between style attention and style returns show that high style-level attention is associated with high concurrent stock returns. Return spreads between the high and low style attention quintiles are significantly positive in four out of five firm-attention quintiles. Findings on the style attention and stock returns relation over longer-horizon are consistent with those documented in Table 4.8. For the 1 month – 1 month strategy, style-level attention negatively forecasts stock returns over the subsequent month. The top style attention quintile underperforms the bottom style attention quintile in all firm-attention groups. Results are qualitatively similar for the 3 month – 3 month, 6 month – 6 month and 12 month – 12 month strategies. Most importantly, the finding that the high style attention quintile underperforms the low style attention quintile is prevalent across all firm-attention groups, suggesting the impact of style-level attention on stock returns is distinctive from the firm-level attention effect documented by previous literature.

The results are comparable in Panel B when $NewsShock$ is used as the proxy for attention. Style-level $NewsShock$ is associated with contemporary stock price increases and price reversal in the long run. This finding is pervasive across all firm-level $NewsShock$ quintiles. Overall, consistent with hypothesis 4b, Table 4.10 suggests that style-level attention has explanatory power for the cross-section of stock returns above and beyond firm-level attention.

[63] Value-weighed returns yield similar results.

Table 4.10 Returns on portfolios double sorted on firm- and style-level attention shocks

This table examines the impact of style-level attention on stock returns controlling for firm-specific attention. At the end of each month, all sample stocks are sorted into quintiles on the firm-level attention in the current month, and over the past 1, 3, 6, and 12 months. Within each firm-level attention quintile, stocks are further sorted into quintiles on the style-level attention over the same formation period, resulting in 25 portfolios. The portfolio return is the equal-weighted average returns across all stocks in the portfolio. The average monthly returns of these portfolios over different holding period are presented in this table. Within each firm-level attention quintile, return difference between the high- and low- style-level attention quintiles is reported (High - Low), along with t-statistics in parentheses. Panel A (B) reports the results using $AnalystShock$ ($NewsShock$) as the proxy for attention shock. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

		Style-level attention						
	Firm-level attention	Low	2	3	4	High	High - Low	t-Statistic
Panel A: Returns on portfolios double sorted on firm- and style-level $AnalystShock$								
	Low	0.0057	0.0085	0.0085	0.0103	0.0124	0.0068***	(3.23)
	2	0.0065	0.0099	0.0120	0.0098	0.0145	0.0080***	(2.93)
Current month	3	0.0070	0.0077	0.0125	0.0094	0.0118	0.0048*	(1.69)
	4	0.0098	0.0104	0.0092	0.0117	0.0134	0.0036	(1.10)
	High	0.0076	0.0069	0.0105	0.0121	0.0118	0.0042*	(1.68)
	Low	0.0154	0.0159	0.0123	0.0132	0.0082	-0.0076***	(-3.27)
	2	0.0137	0.0107	0.0094	0.0098	0.0078	-0.0059**	(-2.23)
1 month – 1 month	3	0.0107	0.0085	0.0105	0.0098	0.0104	-0.0090***	(-3.49)
	4	0.0105	0.0086	0.0093	0.0066	0.0061	-0.0063***	(-2.67)
	High	0.0105	0.0103	0.0061	0.0071	0.0075	-0.0083***	(-3.61)
	Low	0.0133	0.0122	0.0097	0.0085	0.0073	-0.0060***	(-3.58)
	2	0.0108	0.0090	0.0083	0.0080	0.0081	-0.0027*	(-1.93)
3 month – 3 month	3	0.0101	0.0087	0.0089	0.0074	0.0088	-0.0014	(-1.07)
	4	0.0106	0.0082	0.0080	0.0087	0.0078	-0.0028**	(-2.06)
	High	0.0099	0.0078	0.0076	0.0075	0.0064	-0.0035***	(-3.15)
	Low	0.0103	0.0106	0.0112	0.0074	0.0078	-0.0025**	(-2.39)
	2	0.0093	0.0090	0.0085	0.0066	0.0071	-0.0021**	(-2.40)
6 month – 6 month	3	0.0074	0.0076	0.0078	0.0072	0.0052	-0.0022**	(-2.08)
	4	0.0088	0.0075	0.0074	0.0064	0.0068	-0.0020***	(-2.61)
	High	0.0075	0.0070	0.0072	0.0054	0.0048	-0.0027***	(-3.53)
	Low	0.0093	0.0092	0.0101	0.0082	0.0071	-0.0022**	(-2.48)
	2	0.0067	0.0069	0.0077	0.0064	0.0049	-0.0018**	(-2.54)
12 month – 12 month	3	0.0081	0.0065	0.0073	0.0062	0.0055	-0.0026***	(-4.07)
	4	0.0060	0.0068	0.0060	0.0044	0.0047	-0.0013**	(-1.77)
	High	0.0068	0.0055	0.0052	0.0048	0.0050	-0.0017**	(-2.43)

Table 4.10 continued

Panel B: Returns on portfolios double sorted on firm- and style-level $NewsShock$								
	Low	0.0052	0.0072	0.0057	0.0076	0.0081	0.0029*	(1.63)
	2	0.0040	0.0076	0.0037	0.0088	0.0117	0.0077***	(3.37)
Current month	3	0.0043	0.0115	0.0089	0.0106	0.0127	0.0085***	(3.82)
	4	0.0070	0.0098	0.0121	0.0098	0.0114	0.0044**	(1.99)
	High	0.0187	0.0213	0.0227	0.0201	0.0181	-0.0006	(-0.17)
	Low	0.0118	0.0101	0.0084	0.0123	0.0081	-0.0038**	(-2.08)
	2	0.0106	0.0094	0.0108	0.0108	0.0069	-0.0037	(-1.45)
1 month – 1 month	3	0.0086	0.0104	0.0113	0.0093	0.0072	-0.0014	(-0.51)
	4	0.0090	0.0083	0.0085	0.0077	0.0079	-0.0011	(-0.41)
	High	0.0119	0.0128	0.0061	0.0090	0.0044	-0.0075***	(-2.91)
	Low	0.0070	0.0102	0.0076	0.0112	0.0068	-0.0003	(-0.23)
	2	0.0073	0.0091	0.0078	0.0084	0.0069	-0.0004	(-0.30)
3 month – 3 month	3	0.0072	0.0099	0.0085	0.0069	0.0061	-0.0010	(-0.80)
	4	0.0073	0.0094	0.0073	0.0081	0.0043	-0.0030**	(-2.18)
	High	0.0067	0.0080	0.0094	0.0078	0.0064	-0.0003	(-0.17)
	Low	0.0088	0.0094	0.0090	0.0087	0.0072	-0.0016**	(-2.08)
	2	0.0083	0.0083	0.0101	0.0091	0.0071	-0.0013*	(-1.76)
6 month – 6month	3	0.0086	0.0090	0.0099	0.0082	0.0070	-0.0016*	(-1.85)
	4	0.0082	0.0090	0.0080	0.0083	0.0067	-0.0014*	(-1.86)
	High	0.0079	0.0059	0.0067	0.0062	0.0040	-0.0039***	(-3.64)
	Low	0.0089	0.0085	0.0082	0.0081	0.0066	-0.0023***	(-3.69)
	2	0.0080	0.0088	0.0079	0.0082	0.0064	-0.0017***	(-3.33)
12 month – 12 month	3	0.0083	0.0082	0.0082	0.0082	0.0067	-0.0016***	(-2.77)
	4	0.0089	0.0079	0.0071	0.0075	0.0066	-0.0023***	(-4.26)
	High	0.0084	0.0065	0.0070	0.0061	0.0059	-0.0025***	(-4.45)

4.6 Robustness tests

4.6.1 Style returns and investor attention - Value-weighted style returns

In the main analysis, we equally weight each stock in the style to compute the style returns. As a robustness check, we also calculate the value-weighted style returns. Table 4.11 presents the impact of prior style performance on style- and firm-level attention using value-weighted style returns. Consistent with the findings in Table 4.2, Panel A of Table 4.11 shows that style-level attention is positively related to the style abnormal returns over the past 1 month, 3 months, and 6 months. Consistent with the results in Table 4.3, Panel B suggests that prior style performance has significant effect on firm-level attention even after controlling for prior firm-specific performance. For example, in regression (2), coefficients on $|Ret_{s,t-1}|$, $|Ret_{s,t-3,t-2}|$, $|Ret_{s,t-6,t-4}|$, $|Ret_{s,t-12,t-7}|$ are all positive and statistically significant at the 1% level. The findings on the firm attention and prior firm performance are mixed. Thus, our empirical evidence on investors' style chasing behaviour is not sensitive to the weighting method.

4.6.2 Alternative measures of attention shock

As specified in Equations (4.1a) and (4.1b), we measure attention shock as logarithm of monthly attention detrended by subtracting its 12-month moving average. To ensure that our results are not influenced by the length of the detrending window, we reproduce Tables 4.2 and 4.3 using attention shock calculated based on 6-month and 24-month moving windows. The results are displayed in Table 4.12. Regressions (1) and (2) present the results based on the 24-month estimation window, and regressions (3) and (4) present the results based on the 6-month window.

Consistent with the main analysis, results in Panel A support a positive relation between style-level attention and abnormal style performance using alternative measures of attention shock. Similarly, Panel B provide strong evidence for the significant impact of prior style performance on firm-specific attention. Results are comparable to those reported in Table 4.3. Thus, our results are not influenced by the detrending window employed when computing attention shock.

Table 4.11 Impact of prior style returns on style- and firm-level attention - Value-weighted style returns

This table presents the impact of prior style returns on style-level and firm-level attention. Style returns (Ret_s) are calculated as the value-weighted stock returns within the style. Firm-level attention shock is measured by the detrended number of analyst forecast revisions ($AnalystShock$) and detrended news coverage ($NewsShock$) as specified in Equations (4.1a) and (4.1b). Style-level attention shock is calculated by averaging the individual attention shock across stocks within the same style. In Panel A, style-level attention shock is regressed on the lagged absolute style returns ($|Ret_s|$), style average market capitalization ($Size_s$) and book-to-market ratio (BM_s). In Panel B, firm-level attention shock is regressed on the lagged absolute style returns ($|Ret_s|$), lagged absolute firm returns ($|Ret_i|$), firm size ($Size_i$) and firm book-to-market ratio (BM_i). The table reports the regression results with style and month fixed effects. T-statistics based on style-clustered (firm-clustered) standard errors are presented in parentheses below the coefficient estimates in Panel A (Panel B). ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	(1) $AnalystShock_t$	(2) $NewsShock_t$
Panel A: Style-level attention and past style performance		
$Size_{s,t}$	0.1615***	0.1972***
	(3.49)	(3.85)
$BM_{s,t}$	0.0150	0.0269
	(0.69)	(0.78)
$\lvert Ret_{s,t-1}\rvert$	0.4232**	0.5410***
	(2.55)	(3.17)
$\lvert Ret_{s,t-3,t-2}\rvert$	0.3267**	0.4658***
	(2.70)	(3.25)
$\lvert Ret_{s,t-6,t-4}\rvert$	0.1489**	0.0236
	(2.47)	(0.24)
$\lvert Ret_{s,t-12,t-7}\rvert$	-0.0179	-0.1651**
	(-0.30)	(-2.46)
Constant	-1.5978***	-2.1706***
	(-2.95)	(-3.96)
Observations	3,799	3,775
R-squared	0.7697	0.4501
Panel B: Firm-level attention and past style performance		
$Size_{i,t}$	0.0066**	-0.0039***
	(2.56)	(-2.77)
$BM_{i,t}$	-0.0399***	-0.0260***
	(-13.95)	(-11.21)
$\lvert Ret_{s,t-1}\rvert$	0.0540	0.1791***
	(0.75)	(2.65)
$\lvert Ret_{s,t-3,t-2}\rvert$	0.0914**	0.1332***
	(2.16)	(3.10)
$\lvert Ret_{s,t-6,t-4}\rvert$	0.1841***	0.2024***
	(5.43)	(5.78)
$\lvert Ret_{s,t-12,t-7}\rvert$	0.0430*	0.1095***
	(1.90)	(4.54)
$\lvert Ret_{i,t-1}\rvert$	-0.2005***	0.0090
	(-11.43)	(0.78)
$\lvert Ret_{i,t-3,t-2}\rvert$	0.0146*	0.0494***
	(1.81)	(6.87)
$\lvert Ret_{i,t-6,t-4}\rvert$	0.0011	-0.0135**
	(0.20)	(-2.19)
$\lvert Ret_{i,t-12,t-7}\rvert$	0.0227***	-0.0078
	(6.80)	(-1.24)
Constant	0.0224	0.6235**
	(1.03)	(2.39)
Observations	328,564	430,664
R-squared	0.1089	0.0447

Table 4.12 Impact of prior style returns on style- and firm-level attention - Alternative measures of attention shocks

This table presents the impact of prior style returns on style-level and firm-level attention. Firm-level attention shock is measured by the number of analyst forecast revisions in the current month detrended by the average number of analyst forecast revisions over the past 6 or 24 months ($AnalystShock$) and the news coverage in the current month detrend by the average news coverage over the past 6 or 24 months ($NewsShock$). Style-level attention shock is calculated by averaging the individual attention shock across stocks within the same style. Style returns (Ret_s) are calculated as the equal-weighted stock returns within the style. In Panel A, style-level attention shock is regressed on the lagged absolute style returns, style average market capitalization ($Size_s$) and book-to-market ratio (BM_s). In Panel B, firm-level attention shock is regressed on the lagged absolute style returns($|Ret_s|$), lagged absolute firm returns ($|Ret_i|$) , firm size ($Size_i$) and firm book-to-market ratio (BM_i). The table reports the regression results with style and month fixed effects. T-statistics based on style-clustered (firm-clustered) standard errors are presented in parentheses below the coefficient estimates in Panel A (Panel B). ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	24-month detrending window		6-month detrending window	
	(1)	(2)	(3)	(4)
	$AnalystShock_{s,t}$	$NewsShock_{s,t}$	$AnalystShock_{s,t}$	$AnalystShock_{s,t}$
Panel A: Style-level attention and past style performance				
$Size_{s,t}$	0.1551***	0.2158***	0.1626***	0.0775***
	(3.66)	(2.96)	(6.53)	(2.98)
$BM_{s,t}$	0.0256	0.0296	0.0186	0.0115
	(1.07)	(0.71)	(1.51)	(0.64)
$\lvert Ret_{s,t-1}\rvert$	0.4050**	0.6055**	0.2860*	0.2870*
	(2.21)	(2.35)	(1.87)	(2.03)
$\lvert Ret_{s,t-3,t-2}\rvert$	0.3195**	0.6066***	0.1868*	0.2511**
	(2.25)	(3.57)	(1.92)	(2.32)
$\lvert Ret_{s,t-6,t-4}\rvert$	0.2262**	0.2501**	0.0173	-0.0676
	(2.22)	(2.23)	(0.22)	(-0.89)
$\lvert Ret_{s.t-12.t-7}\rvert$	0.0783	-0.0706	-0.0240	-0.0131
	(1.68)	(-0.77)	(-0.49)	(-0.36)
Constant	-1.5327***	-2.4367***	-1.5144***	-0.9115***
	(-3.12)	(-3.11)	(-5.35)	(-3.07)
Observations	3,799	3,775	3,799	3,775
R-squared	0.7685	0.5006	0.7634	0.4968

Table 4.12 continued

	24-month detrending window		6-month detrending window	
	(1)	(2)	(3)	(4)
	$AnalystShock_{i,t}$	$NewsShock_{i,t}$	$AnalystShock_{i,t}$	$AnalystShock_{i,t}$
Panel B: Firm-level attention and past style performance				
$Size_{i,t}$	0.0016	-0.0185***	0.0043**	-0.0041***
	(0.61)	(-11.94)	(1.99)	(-3.98)
$BM_{i,t}$	-0.0424***	-0.0331***	-0.0278***	-0.0151***
	(-13.60)	(-13.49)	(-12.77)	(-9.10)
$\lvert Ret_{s,t-1}\rvert$	-0.0426	0.1329**	0.0830	0.2193***
	(-0.59)	(2.00)	(1.12)	(3.27)
$\lvert Ret_{s,t-3,t-2}\rvert$	0.1034**	0.0955**	0.0982**	0.1449***
	(2.41)	(2.37)	(2.27)	(3.43)
$\lvert Ret_{s,t-6,t-4}\rvert$	0.1699***	0.1852***	0.1386***	0.1461***
	(5.71)	(6.26)	(4.47)	(4.68)
$\lvert Ret_{s,t-12,t-7}\rvert$	0.0408**	0.1172***	0.0301*	0.0628***
	(2.13)	(6.28)	(1.67)	(3.38)
$\lvert Ret_{i,t-1}\rvert$	-0.1913***	0.0067	-0.2293***	-0.0499***
	(-15.63)	(0.63)	(-11.87)	(-4.51)
$\lvert Ret_{i,t-3,t-2}\rvert$	0.0088	0.0577***	0.0136*	0.0174*
	(1.18)	(8.36)	(1.78)	(1.88)
$\lvert Ret_{i,t-6,t-4}\rvert$	-0.0120**	0.0017	-0.0120**	-0.0608***
	(-2.20)	(0.32)	(-2.08)	(-6.63)
$\lvert Ret_{i,t-12,t-7}\rvert$	0.0189***	0.0042	0.0275***	0.0147***
	(5.46)	(1.60)	(8.07)	(3.11)
Constant	0.1229***	-0.0881***	0.0293	-0.0883***
	(5.76)	(-3.72)	(1.50)	(-6.34)
Observations	323,692	427,320	336,928	445,543
R-squared	0.1168	0.0555	0.0958	0.0388

4.6.3 Alternative measures of comovement

To account for the market component of investor attention when constructing attention comovement, we also regress firm specific attention on style attention and market attention. For each firm, we estimate the following regression using weekly attention measures over a 12-month window to obtain R^2:

$$Attention_{i,w} = \alpha_i + \beta_1 Attention_{Style,w} + \beta_2 Attention_{M,w} + \varepsilon_{i,w} \qquad (4.12)$$

where $Attention_{i,w}$ is the firm-specific attention of stock i in week w.$Attention_{Style,w}$ is the style attention, computed as the equally weighted attention for the style in week w (excluding stock i). $Attention_{M,w}$ is the equally weighted market attention in week w (excluding stock i). Attention comovement is estimated by taking the log transformation of the regression R^2 as in Equation (4.3). Following the same method, we construct return comovement (turnover comovement) by regressing individual stock returns (turnover) on style turnover and market turnover.

We rerun Equation (4.10) using these alternative comovement measures, and present the results in Table 4.13. The findings are consistent with those in Table 4.5. Attention comovement is positively associated with return comovement (Panel A) and turnover comovement (Panel B). Therefore, the information flow explanation for the within-style price comovement holds using alternative comovement measures.

.

Table 4.13 Return comovement , turnover comovement and attention comovement - An alternative measure of comovement

This table presents how attention comovement is associated with return comovement and turnover comovement. Panel A reports the contemporaneous relation between attention comovement and return comovement and Panel B reports the relation between attention comovement and turnover comovement. $RetComove_{i,t}$ represents the return comovement for stock *i* in year *t*, measured using the R^2 from regressing weekly returns of stock *i* on equal-weighted style returns and market returns, and taking the log transformation. $TOComove_{i,t}$ represents the turnover comovement for stock *i* in year *t*, measured using the R^2 from regressing weekly turnover of stock *i* on equal-weighted style turnover and market turnover, and taking the log transformation. $AnalystComove$ and $NewsComove$ are measures of attention comovement based on the number of analyst forecast revisions and news coverage, respectively. All control variables are defined in Equation (4.10). The table reports the panel regression results with style and year fixed effects. T-statistics based on firm-clustered standard errors are presented in parentheses below the coefficient estimates. ***, **, and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

	Panel A: $RetComove_{i,t}$		Panel B: $TOComove_{i,t}$	
	(1)	(2)	(3)	(4)
$AnalystComove_{i,t}$	0.0394***		0.0824***	
	(6.56)		(16.26)	
$NewsComove_{i,t}$		0.0160***		0.0736***
		(2.80)		(14.58)
$ROA_{i,t}$	-0.1395**	-0.2463***	0.0948*	-0.0180
	(-2.32)	(-4.06)	(1.83)	(-0.35)
$Size_{i,t}$	0.4028***	0.5300***	0.2731***	0.3524***
	(21.10)	(28.86)	(19.91)	(25.86)
$BM_{i,t}$	0.0845***	0.1070***	0.0090	0.0061
	(5.57)	(6.91)	(0.65)	(0.44)
$SalesGrowth_{i,t}$	-0.0411**	-0.0429**	-0.0578***	-0.0497**
	(-2.07)	(-2.04)	(-3.00)	(-2.45)
$IO_{i,t}$	0.0831***	0.1048***	0.0272***	0.0331***
	(7.53)	(9.81)	(3.09)	(4.01)
$Analyst_{i,t}$	-0.0291	0.0309	0.0090	0.0207
	(-1.34)	(1.48)	(0.51)	(1.22)
$StdROA_{i,t}$	-0.0972	-0.0786	0.0310	-0.0042
	(-1.20)	(-1.01)	(0.43)	(-0.06)
$AbsRet_{i,t}$	-0.2407***	-0.2646***	0.1517***	0.1706***
	(-10.88)	(-11.82)	(7.22)	(8.11)
$Turnover_{i,t}$	-0.4600***	-0.1710**	-0.5664***	-0.3437***
	(-6.75)	(-2.51)	(-9.57)	(-5.91)
$Price_{i,t}$	-0.0455***	-0.0902***	-0.0180	-0.0523***
	(-2.88)	(-5.73)	(-1.49)	(-4.31)
$ROACom_{i,t}$	0.0385***	0.0364***	0.0225***	0.0191***
	(8.52)	(7.81)	(5.38)	(4.39)
Constant	-3.6104***	-4.2350***	-3.7341***	-4.0250***
	(-38.72)	(-46.74)	(-50.07)	(-54.09)
Style fixed-effects	Yes	Yes	Yes	Yes
Year fixed-effects	Yes	Yes	Yes	Yes
Observations	34,994	33,848	36,649	35,984
R-squared	0.2706	0.3284	0.2484	0.2643

4.7 Conclusion

This chapter links investor attention to style investing by studying how investors allocate attention across investment styles and its implications for style return predictability. Barberis and Shleifer (2003) predicts that investors categorize assets into styles and shift funds across styles based on their past performance. Building on this prediction, we first investigate whether investor attention is affected by prior style performance. Style-level analysis suggests that investors allocate more attention to styles that demonstrate extreme past performance. At the firm-level, the results show that firm-specific attention is significantly affected by the prior style performance after controlling for firm-level performance. Therefore, our findings provide supportive evidence for Barberis and Shleifer's (2003) style chasing model.

We then explore the asset-pricing implications of style-level attention. First, we investigate whether attention comovement helps explain the within-style return comovement. We conjecture that when investors systematically seek out information for a certain type of stocks, their attention comoves. As a result, capital allocations may similarly comove, generating comovement in returns. Consistent with this conjecture, we document stronger attention comovement and return comovement among high-attention styles. A stock's return comovement is also positively associated with its attention comovement after controlling for comovement in fundamentals. Furthermore, after style reclassficiation, a stock's attention comovement and return comovement with the new (old) style rises (falls). Our findings provide an information flow explanation for the within-style return comovement.

Second, we investigate whether style-level attention helps explain the variation in cross-sectional stock returns. Consistent with Barber and Odean's (2008) price pressure hypothesis, we show that style-level attention is associated with a contemporaneous price increase and a subsequent reversal in both style and individual stock returns. Further, we show that style-level attention contributes to autocorrelation in style-returns documented by prior literature.

Chapter 5

Conclusion

This thesis empirically tests how the process of information affects asset prices in three specific settings: (1) cross-listed stocks, (2) intra-industry, and (3) style investing. Empirical analyses are presented in Chapters 2 to 4. The findings have important implications in addressing return patterns that seem anomalous to traditional finance theories on comovement in asset prices. This chapter concludes the thesis by summarizing the empirical analysis conducted in each chapter, reiterating the contributions, and providing directions for future research.

5.1 Summary of the empirical findings

Chapter 2 investigates the driving forces of investor attention, and explores the asset-pricing implication of correlated attention for fundamentally linked securities. Using trading volume shock as a proxy for investor attention, the results show that up to a quarter of the variation in firm-specific attention can be explained by attention on the within-pair counterpart. Further analysis documents an interesting pattern in attention comovement across regions and time periods. Canadian and European firms cross listed in the US market exhibit the strongest attention comovement, while firms with home markets domiciled in Asia-Pacific region exhibit the lowest attention comovement. Also, there is an upward trend in attention comovement over our sample period.

The existence of attention comovement calls for an investigation of its determinants. In this thesis, we examine how attention comovement is related to the information environment, information shocks, stock market integration, and aggregate investor attention. The results show that attention comovement is positively related to a firm's information transparency, the frequency of information shocks, and aggregate attention in the US market. Thus, we provide supportive evidence for both information- and socially-driven explanations of investor attention.

Finally, we investigate how attention comovement affects deviations from price parity for cross-listed stock pairs. Building upon Hong and Stein's (1999) gradual information diffusion model, we argue that high attention comovement is an indication that information is impounded into stock pairs at a similar speed. Therefore, high attention comovement is expected to be associated with less price disparity. Consistent with this hypothesis, we document a significant negative relation between attention comovement and price disparity. The negative relation remains robust after controlling for the impediments to arbitrage identified by previous studies. Hence, attention comovement provides incremental explanatory power for price deviations in cross-listed stocks.

Chapter 3 investigates how information is transmitted across firms within the same industry and its implication for return comovement. The chapter is built on Veldkamp's (2006) model of a competitive information market. Motivated by Veldkamp's (2006) theoretical prediction, we study information flows between industry leaders (bellwether firms) and peer firms. Empirical results indicate a unidirectional information spillover from bellwether firms to their industry peers. News of bellwether firms significantly affects industry peers' stock prices, trading activity and analyst forecasts. Using a firm's partial correlation in news (with other firms in the same industry) to gauge the firm's contribution in explaining news of other firms, we find that bellwether firms exhibit a higher partial correlation in news.

We then explore the asset-pricing implication of the observed intra-industry information spillover. Veldkamp (2006) suggests that return comovement arises from investors using a common subset of information for asset valuation. It is plausible that when a stock's news is incorporated into the prices of many other stocks, its stock price comoves with the market. For this reason, we conjecture that firms with more contributing news exhibit stronger return comovement. In agreement with this conjecture, we show that a firm's news partial correlation is positively associated with its return comovement. Also, firms with more contributing news exhibit a lower degree of mispricing. Thus, our findings provide evidence that favours a positive relation between return comovement and price informativeness.

Chapter 4 examines how investors allocate attention across different style portfolios, and whether style-level attention contributes to style-related return patterns documented in the earlier studies. This chapter is built on Barberis and Shleifer's (2003) theoretical model on style investing, which argues that investors invest at the style level rather than at the individual security level, and they switch funds across style portfolios based on their relative past performance. The results suggest that past style performance plays a significant role in determining both style-level and firm-specific attention. Importantly, investor attention to a specific firm is found to be primarily driven by style-level performance rather than the firm-level performance. Overall, our findings validate Barberis and Shleifer's (2003) style performance chasing assumption.

The chapter further investigates whether investor attention helps explain the puzzling style-related return predictability. When a stock's firm-level attention is largely explained by the aggregate attention paid to its style, its price should incorporate a large amount of style-level information, leading to a high level of return comovement with the style. This conjecture is confirmed in the empirical analyses - there is a significant positive relation between attention comovement and return comovement. Also, when a stock is reclassified into a new style, both its attention and returns covary more with that style. This finding supports the category investment argument that investors view firms in the context of style categories.

Chapter 4 ends with an investigation of whether style-level attention contributes to autocorrelation in style returns. Barberis and Shleifer (2003) suggest that style chasing temporarily pushes prices away from fundamentals, leading to short-term price momentum and subsequent price reversals. As investors are more likely to chase the styles that attract their attention, this pattern is expected to be stronger among more attention-grabbing styles. This conjecture is supported by the empirical results that, more attention-grabbing style portfolios are associated with stronger short-term price momentum and long-term price reversals.

5.2 Contributions

The first major purpose of the thesis is to better understand the underlying drivers of investor attention. The investigation of cross-listed stocks allows us to directly examine different driving forces of investor attention, a research issue that has not been explicitly addressed in the literature. The findings provide empirical support for both rational and behavioural views on investor attention. Furthermore, the finding that correlated attention has explanatory power for deviations from price parity in cross-listed stock pairs contributes to the literature on ADR mispricing. Deviations from price parity in cross-listed stocks are widely documented. We show that attention comovement has explanatory power for price deviations after controlling for impediments to arbitrage identified by prior studies. Thus, we provide an alternative explanation for the existence of price deviations from a channel related to information flows. Our finding also sheds light on some other long-standing empirical puzzles, in which mispricing exists between similar assets.

Second, this thesis is the first to establish a direct link between intra-industry information production and return comovement. Veldkamp (2006) suggests that information production affects return comovement. In the presence of costly information, investors purchase signals that can predict the value of many assets. The use of a common subset of information leads to common movement in asset returns. Our finding that investors use news about bellwether firms to update the value of industry peers validates Veldkamp's (2006) theoretical framework. In accordance with

Veldkamp (2006), we show that firms whose news is more informative to industry peers exhibit stronger return comovement and less mispricing. The finding that return comovement is positively associated with price informativeness adds to the long-standing debate on the information implication of return comovement.

Finally, this thesis contributes to the literature by linking investor attention to style investing. The existing style investing literature largely focuses on the asset-pricing outcomes without necessarily investigating the information flows leading to these outcomes. We fill this gap by directly testing the information flow assumptions underlying Barberis and Shleifer's (2003) style investing model. The results suggest that style is an important factor in driving the cross-sectional variation in stock returns and attention. The finding improves our understanding of how investors pay attention to individual stocks and the associated asset-pricing implications.

5.3 Directions for future research

This thesis examines the asset-pricing consequences of correlated information flows by focusing on three specific settings: (1) cross-listed stock pairs, (2) intra-industry news spillover, and (3) investment styles. This subsection provides guidance for two natural extensions of this study which may be useful for future research.

First, the analysis of the implication of correlated attention for deviations from price parity can be extended to other fundamentally linked securities. Apart from ADRs' discount or premium, prior studies also document a number of other violations of the law of one price, such as price deviations in Siamese Twin stocks and dual share classes, and mispricing associated with equity carve-outs and spin-offs. Future research can further investigate how the processing of information affects mispricing in those fundamentally linked assets.

Second, it may be interesting to extend the analyses in this thesis to other international markets. For example, Chapter 2 focuses on foreign stocks that are listed in the US stock exchanges. However, there has been a significant drop in the US cross-listings following the passage of the Sarbanes-Oxley Act (SOX) in 2002. In contrast, the total number of depositary receipt (DR) programs has experienced steady growth over the same period, with more firms choosing European and Asian stock markets as the destination of cross-listings. Future studies may also bring the non-US markets into consideration. The heterogeneous institutional features and different market characteristics across countries can shed light on how macro-economic environment alters the asset-pricing consequences of correlated information flows.

References

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

Andrei, D., & Hasler, M. (2014). Investor attention and stock market volatility. *Review of Financial Studies*, 28(1), 33-72.

Anton, M., & Polk, C. (2014). Connected stocks. *Journal of Finance, 69*(3), 1099-1127.

Asness, C. S., Liew, J. M., & Stevens, R. L. (1997). Parallels between the cross-sectional predictability of stock and country returns. *Journal of Portfolio Management*, *23*(3), 79.

Bacidore, J. M., & Sofianos, G. (2002). Liquidity provision and specialist trading in NYSE-listed non-US stocks. *Journal of Financial Economics*, *63*(1), 133-158.

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, *21*(2), 129-152.

Baker, M., Wurgler, J., & Yuan, Y. (2012). Global, local, and contagious investor sentiment. *Journal of Financial Economics*, *104*(2), 272-287.

Bali, T. G., Peng, L., Shen, Y., & Tang, Y. (2013). Liquidity shocks and stock market reactions. *The Review of Financial Studies*, *27*(5), 1434-1485.

Bank, M., Larch, M., & Peter, G. (2011). Google search volume and its influence on liquidity and returns of German stocks. *Financial Markets and Portfolio Management*, *25*(3), 239.

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, *21*(2), 785-818.

Barber, B. M., Lehavy, R., McNichols, M., & Trueman, B. (2006). Buys, holds, and sells: The distribution of investment banks' stock ratings and the implications for the profitability of analysts' recommendations. *Journal of accounting and Economics*, *41*(1-2), 87-117.

Barberis, N., & Shleifer, A. (2003). Style investing. *Journal of Financial Economics*, *68*(2), 161-199.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, *49*(3), 307-343.

Barberis, N., Shleifer, A., & Wurgler, J. (2005). Comovement. *Journal of Financial Economics*, *75*(2), 283-317.

Bartram, S. M., Brown, G., & Stulz, R. M. (2012). Why are US stocks more volatile?. *Journal of Finance*, *67*(4), 1329-1370.

Bekaert, G., Hodrick, R. J., & Zhang, X. (2009). International stock return comovements. *Journal of Finance*, *64*(6), 2591-2626.

Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *Review of Financial Studies*, *30*(9), 3009-3047.

Berben, R. P., & Jansen, W. J. (2005). Comovement in international equity markets: A sectoral view. *Journal of International Money and Finance*, *24*(5), 832-857.

Bianconi, M., Chen, R., & Yoshino, J. A. (2013). Firm value, the Sarbanes-Oxley Act and cross-listing in the US, Germany and Hong Kong destinations. *North American Journal of Economics and Finance*, *24*, 25-44.

Bodurtha, J. N., Kim, D. S., & Lee, C. M. (1995). Closed-end country funds and US market sentiment. *Review of Financial Studies*, *8*(3), 879-918.

Box, T., & Shang, D. (2018). Information driven stock price comovement. *Working Paper.* University of Mississippi and Utah State University.

Boyer, B. H. (2011). Style-Related Comovement: Fundamentals or labels?. *Journal of Finance*, *66*(1), 307-332.

Brandt, M. W., Brav, A., Graham, J. R., & Kumar, A. (2009). The idiosyncratic volatility puzzle: Time trend or speculative episodes?. *Review of Financial Studies*, *23*(2), 863-899.

Brockman, P., Liebenberg, I., & Schutte, M. (2010). Comovement, information production, and the business cycle. *Journal of Financial Economics*, *97*(1), 107-129.

Brooks, R., & Del Negro, M. (2004). The rise in comovement across national stock markets: market integration or IT bubble? *Journal of Empirical Finance*, *11*(5), 659-680.

Bushman, R. M., Piotroski, J. D., & Smith, A. J. (2004). What determines corporate transparency?. *Journal of Accounting Research*, *42*(2), 207-252.

Campbell, J. Y., Lettau, M., Malkiel, B. G., & Xu, Y. (2001). Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance*, *56*(1), 1-43.

Cao, J., Chordia, T., & Lin, C. (2016). Alliances and return predictability. *Journal of Financial and Quantitative Analysis*, *51*(5), 1689-1717.

Chan, K., & Chan, Y. C. (2014). Price informativeness and stock return synchronicity: Evidence from the pricing of seasoned equity offerings. *Journal of Financial Economics*, *114*(1), 36-53.

Chan, K., Hameed, A., & Kang, W. (2013). Stock price synchronicity and liquidity. *Journal of Financial Markets*, *16*(3), 414-438.

Chemmanur, T., & Yan, A. (2009). Product market advertising and new equity issues. *Journal of Financial Economics*, *92*(1), 40-65.

Chen, C., Huang, A. G., & Jha, R. (2012). Idiosyncratic return volatility and the information quality underlying managerial discretion. *Journal of Financial and Quantitative Analysis*, *47*(4), 873-899.

Chen, H. J., Chen, S. J., & Li, F. (2012). Empirical investigation of an equity pairs trading strategy. *Working Paper*. Tsinghua University, University of British Columbia and Shanghai Jiaotong University.

Chen, H. L., & De Bondt, W. (2004). Style momentum within the S&P-500 index. *Journal of Empirical Finance*, *11*(4), 483-507.

Chen, H., & Choi, P. M. S. (2012). Does information vault Niagara falls? Cross-listed trading in New York and Toronto. *Journal of Empirical Finance*, *19*(2), 175-199.

Chen, Q., Goldstein, I., & Jiang, W. (2006). Price informativeness and investment sensitivity to stock price. *The Review of Financial Studies*, *20*(3), 619-650.

Chowdhry, B., & Nanda, V. (1991). Multimarket trading and market liquidity. *Review of Financial Studies*, *4*(3), 483-511.

Claessens, S., & Yafeh, Y. (2012). Comovement of newly added stocks with national market indices: Evidence from around the world. *Review of Finance*, *17*(1), 203-227.

Cohen, L., & Frazzini, A. (2008). Economic links and predictable returns. *Journal of Finance*, *63*(4), 1977-2011.

Cohen, L., & Lou, D. (2012). Complicated firms. *Journal of Financial Economics*, *104*(2), 383-400.

Cohen, L., Frazzini, A., & Malloy, C. (2008). The small world of investing: Board connections and mutual fund returns. *Journal of Political Economy*, *116*(5), 951-979.

Collins, D. W., Kothari, S. P., & Rayburn, J. D. (1987). Firm size and the information content of prices with respect to earnings. *Journal of Accounting and Economics*, *9*(2), 111-138.

Conrad, J., Cornell, B., & Landsman, W. R. (2002). When is bad news really bad news?. *Journal of Finance*, *57*(6), 2507-2532.

Cooper, M. J., Gulen, H., & Rau, P. R. (2005). Changing names with style: Mutual fund name changes and their effects on fund flows. *Journal of Finance*, *60*(6), 2825-2858.

Cooper, M. J., Gutierrez Jr, R. C., & Hameed, A. (2004). Market states and momentum. *Journal of Finance*, *59*(3), 1345-1365.

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *Journal of Finance,* 66(5), 1461-1499.

Dang, T. L., Moshirian, F., & Zhang, B. (2015). Commonality in news around the world. *Journal of Financial Economics*, *116*(1), 82-110.

Dasgupta, S., Gan, J., & Gao, N. (2010). Transparency, price informativeness, and stock return synchronicity: Theory and evidence. *Journal of Financial and Quantitative Analysis*, *45*(5), 1189-1220.

Davis, J. C., & Henderson, J. V. (2008). The agglomeration of headquarters. *Regional Science and Urban Economics*, *38*(5), 445-460.

De Jong, A., Rosenthal, L., & Van Dijk, M. A. (2009). The risk and return of arbitrage in dual-listed companies. *Review of Finance*, *13*(3), 495-520.

De Jong, F., & De Roon, F. A. (2005). Time-varying market integration and expected returns in emerging markets. *Journal of Financial Economics*, *78*(3), 583-613.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, *98*(4), 703-738.

DellaVigna, S., & Pollet, J. M. (2009). Investor inattention and Friday earnings announcements. *Journal of Finance*, *64*(2), 709-749.

Diamond, D. W., & Verrecchia, R. E. (1981). Information aggregation in a noisy rational expectations economy. *Journal of Financial Economics*, *9*(3), 221-235.

Ding, F., Chen, M., & Wu, Z. (2014). Do institutional investors use earnings forecasts from financial analysts? Evidence from China's stock market. *Emerging Markets Finance and Trade, 50*(sup1), 134-147.

Doidge, C., Karolyi, G. A., & Stulz, R. M. (2004). Why are foreign firms listed in the US worth more? *Journal of Financial Economics*, *71*(2), 205-238.

Doidge, C., Karolyi, G. A., & Stulz, R. M. (2010). Why do foreign firms leave US equity markets? *Journal of Finance*, *65*(4), 1507-1553.

Dougal, C., Engelberg, J., Garcia, D., & Parsons, C. A. (2012). Journalists and the stock market. *Review of Financial Studies*, *25*(3), 639-679.

Drake, M. S., Jennings, J., Roulstone, D. T., & Thornock, J. R. (2017). The comovement of investor attention. *Management Science, 63*(9), 2847-2867.

Drake, M. S., Roulstone, D. T., & Thornock, J. R. (2012). Investor information demand: Evidence from Google searches around earnings announcements. *Journal of Accounting Research*, *50*(4), 1001-1040.

Drake, M. S., Roulstone, D. T., & Thornock, J. R. (2015). The determinants and consequences of information acquisition via EDGAR. *Contemporary Accounting Research*, *32*(3), 1128-1161.

Durnev, A., Morck, R., & Yeung, B. (2004). Value-enhancing capital budgeting and firm-specific stock return variation. *Journal of Finance*, *59*(1), 65-105.

Durnev, A., Morck, R., Yeung, B., & Zarowin, P. (2003). Does greater firm-specific return variation mean more or less informed stock pricing?. *Journal of Accounting Research*, *41*(5), 797-836.

Engelberg, J. E., & Parsons, C. A. (2011). The causal impact of media in financial markets. *Journal of Finance*, *66*(1), 67-97.

Engelberg, J., Gao, P., & Jagannathan, R. (2009). An anatomy of pairs trading: The role of idiosyncratic news, common Information and liquidity. *Working Paper.* University of North Carolina, University of Notre Dame and Northwestern University.

Engelberg, J., Ozoguz, A., & Wang, S. (2018). Know thy neighbour: Industry clusters, information spillovers, and market efficiency. *Journal of Financial and Quantitative Analysis*, *53*(5), 1937-1961.

Eun, C. S., & Sabherwal, S. (2003). Cross-border listings and price discovery: Evidence from US-Listed Canadian stocks. *Journal of Finance*, *58*(2), 549-575.

Eun, C. S., & Shim, S. (1989). International transmission of stock market movements. *Journal of Financial and Quantitative Analysis*, *24*(2), 241-256.

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

Fama, E. F., & French, K. R. (1995). Size and book-to-market factors in earnings and returns. *Journal of Finance*, *50*(1), 131-155.

Fama, E. F., & French, K. R. (1997). Industry costs of equity. *Journal of Financial Economics*, *43*(2), 153-193.

Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *Journal of Finance*, *64*(5), 2023-2052.

Foerster, S. R., & Karolyi, G. A. (1999). The effects of market segmentation and investor recognition on asset prices: Evidence from foreign stocks listing in the United States. *Journal of Finance*, *54*(3), 981-1013.

Frankel, R., Kothari, S. P., & Weber, J. (2006). Determinants of the informativeness of analyst research. *Journal of Accounting and Economics*, *41*(1-2), 29-54.

Freeman, R. N. (1987). The association between accounting earnings and security returns for large and small firms. *Journal of Accounting and Economics*, *9*(2), 195-228.

Froot, K. A., & Dabora, E. M. (1999). How are stock prices affected by the location of trade?. *Journal of Financial Economics*, *53*(2), 189-216.

Froot, K., & Teo, M. (2008). Style investing and institutional investors. *Journal of Financial and Quantitative Analysis*, *43*(4), 883-906.

Gagnon, L., & Karolyi, G. A. (2010). Multi-market trading and arbitrage. *Journal of Financial Economics*, *97*(1), 53-80.

Gallemore, J., & Labro, E. (2015). The importance of the internal information environment for tax avoidance. *Journal of Accounting and Economics*, *60*(1), 149-167.

Gervais, S., Kaniel, R., & Mingelgrin, D. H. (2001). The high-volume return premium. *Journal of Finance*, *56*(3), 877-919.

Goetzmann, W. N., Li, L., & Rouwenhorst, K. G. (2005). Long-term global market correlations. *Journal of Business*, *78*(1).

Grammig, J., Melvin, M., & Schlag, C. (2005). Internationally cross-listed stock prices during overlapping trading hours: price discovery and exchange rate effects. *Journal of Empirical Finance*, *12*(1), 139-164.

Green, T. C., & Hwang, B. H. (2009). Price-based return comovement. *Journal of Financial Economics*, *93*(1), 37-50.

Greenwood, R. (2007). Excess comovement of stock returns: Evidence from cross-sectional variation in Nikkei 225 weights. *Review of Financial Studies*, *21*(3), 1153-1186.

Greenwood, R. M., & Sosner, N. (2007). Trading patterns and excess comovement of stock returns. *Financial Analysts Journal*, *63*(5), 69-81.

Griffin, J. M., Hirschey, N. H., & Kelly, P. J. (2011). How important is the financial media in global markets?. *Review of Financial Studies*, *24*(12), 3941-3992.

Griffin, J. M., Kelly, P. J., & Nardari, F. (2010). Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets. *Review of Financial Studies*, *23*(8), 3225-3277.

Grinblatt, M., & Keloharju, M. (2001). How distance, language, and culture influence stockholdings and trades. *Journal of Finance*, *56*(3), 1053-1073.

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

Grossmann, A., Ozuna, T., & Simpson, M. W. (2007). ADR mispricing: Do costly arbitrage and consumer sentiment explain the price deviation? *Journal of International Financial Markets, Institutions and Money*, *17*(4), 361-371.

Grullon, G., Kanatas, G., & Weston, J. P. (2004). Advertising, breadth of ownership, and liquidity. *Review of Financial Studies*, *17*(2), 439-461.

Grullon, G., Underwood, S., & Weston, J. P. (2014). Comovement and investment banking networks. *Journal of Financial Economics*, *113*(1), 73-89.

Gu, L. (2016). Product market competition, R&D investment, and stock returns. *Journal of Financial Economics*, *119*(2), 441-455.

Hameed, A., & Xie, J. (2019). Preference for dividends and return comovement. *Journal of Financial Economics*, *132*(1), 103-125.

Hameed, A., Morck, R., Shen, J., & Yeung, B. (2015). Information, analysts, and stock return comovement. *Review of Financial Studies*, *28*(11), 3153-3187.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *Journal of Finance*, *50*(4), 1175-1199.

Hendershott, T., Livdan, D., & Schürhoff, N. (2015). Are institutions informed about news?. *Journal of Financial Economics*, *117*(2), 249-287.

Heston, S. L., & Sinha, N. R. (2014). News versus sentiment: Comparing textual processing approaches for predicting stock returns. *Working Paper.* University of Maryland.

Hirshleifer D (2015) Behavioral finance. *Annual Review of Financial Economics*. 7:133–159.

Hirshleifer, D., & Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics*, *36*(1-3), 337-386.

Hirshleifer, D., & Teoh, S. H. (2009). Thought and behaviour contagion in capital markets. In T. Hens and K. Schenk-Hoppe (Eds.), *Handbook of Financial Markets: Dynamics and Evolution*, Handbooks in Finance, Chapter 1, pp. 1-46. Amsterdam, The Netherlands: North-Holland.

Hirshleifer, D., Lim, S. S., & Teoh, S. H. (2009). Driven to distraction: Extraneous events and underreaction to earnings news. *Journal of Finance*, *64*(5), 2289-2325.

Hoberg, G., & Phillips, G. (2010). Product market synergies and competition in mergers and acquisitions: A text-based analysis. *Review of Financial Studies*, *23*(10), 3773-3811.

Hoberg, G., & Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, *124*(5), 1423-1465.

Höchstötter, M., Meyer, S., Riordan, R., & Storkenmaier, A. (2014). International stock market comovement and news. *Journal of Financial Research*, *37*(4), 519-542.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance*, *54*(6), 2143-2184.

Hong, H., Lim, T., & Stein, J. C. (2000). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance*, *55*(1), 265-295.

Hong, H., Torous, W., & Valkanov, R. (2007). Do industries lead stock markets? *Journal of Financial Economics*, *83*(2), 367-396

Hou, K. (2007). Industry information diffusion and the lead-lag effect in stock returns. *Review of Financial Studies*, *20*(4), 1113-1138.

Hou, K., & Robinson, D. T. (2006). Industry concentration and average stock returns. *Journal of Finance*, *61*(4), 1927-1956.

Hou, K., Peng, L., & Xiong, W. (2013). Is R2 a measure of market inefficiency. *Working Paper.* Ohio State University, City University of New York, Princeton University and National Bureau of Economic Research.

Hou, K., Peng, L., & Xiong, W.(2009). A tale of two anomalies: The implications of investor attention for price and earnings momentum. *Working Paper*. Ohio State University, City University of New York and Princeton University.

Huang, S., Huang, Y., & Lin, T. C. (2019). Attention allocation and return co-movement: Evidence from repeated natural experiments. *Journal of Financial Economics*, *132*(2), 369-383.

Huang, X. (2015). Thinking outside the borders: investors' underreaction to foreign operations information. *Review of Financial Studies*, *28*(11), 3109-3152.

Hutton, A. P., Marcus, A. J., & Tehranian, H. (2009). Opaque financial reports, R2, and crash risk. *Journal of Financial Economics*, *94*(1), 67-86.

In, F., Kim, S., Yoon, J. H., & Viney, C. (2001). Dynamic interdependence and volatility transmission of Asian stock markets: Evidence from the Asian crisis. *International Review of Financial Analysis*, *10*(1), 87-96.

Irvine, P. J., & Pontiff, J. (2008). Idiosyncratic return volatility, cash flows, and product market competition. *Review of Financial Studies*, *22*(3), 1149-1177.

Israelsen, R. D. (2016). Does common analyst coverage explain excess comovement?. *Journal of Financial and Quantitative Analysis*, *51*(4), 1193-1229.

Jacobs, H., & Weber, M. (2015). On the determinants of pairs trading profitability. *Journal of Financial Markets*, *23*, 75-97.

Jame, R., & Tong, Q. (2014). Industry-based style investing. *Journal of Financial Markets*, *19*, 110-130.

Jin, L., & Myers, S. C. (2006). R2 around the world: New theory and new tests. *Journal of Financial Economics*, *79*(2), 257-292.

Joseph, K., Wintoki, M. B., & Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. *International Journal of Forecasting*, *27*(4), 1116-1127.

Kahneman, D. (1973). Attention and effort. Englewood Cliffs, NJ: Prentice-Hall.

Kallberg, J., & Pasquariello, P. (2008). Time-series and cross-sectional excess comovement in stock indexes. *Journal of Empirical Finance*, *15*(3), 481-502.

Karlsson, N., Loewenstein, G., & Seppi, D. (2009). The ostrich effect: Selective attention to information. *Journal of Risk and Uncertainty*, *38*(2), 95-115.

Kaniel, R., Ozoguz, A., & Starks, L. (2012). The high volume return premium: Cross-country evidence. *Journal of Financial Economics*, *103*(2), 255-279.

Kelly, P. J. (2014). Information efficiency and firm-specific return variation. *Quarterly Journal of Finance*, *4*(04), 1450018.

Kim, Y. H., & Meschke, F. (2014). CEO interviews on CNBC. *Working Paper.* Nanyang Business School and University of Kansas.

Kizys, R., & Pierdzioch, C. (2009). Changes in the international comovement of stock returns and asymmetric macroeconomic shocks. *Journal of International Financial Markets, Institutions and Money*, *19*(2), 289-305.

Klibanoff, P., Lamont, O., & Wizman, T. A. (1998). Investor reaction to salient news in closed-end country funds. *Journal of Finance*, *53*(2), 673-699.

Kumar, A. (2009). Dynamic style preferences of individual investors and stock returns. *Journal of Financial and Quantitative Analysis*, *44*(3), 607-640.

Kumar, A., & Lee, C. M. (2006). Retail investor sentiment and return comovements. *Journal of Finance*, *61*(5), 2451-2486.

Kumar, A., Page, J. K., & Spalt, O. G. (2012). Investor sentiment and return comovements: Evidence from stock splits and headquarters changes. *Review of Finance*, *17*(3), 921-953.

Kumar, A., Page, J. K., & Spalt, O. G. (2016). Gambling and comovement. *Journal of Financial and Quantitative Analysis*, *51*(1), 85-111.

La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2006). What works in securities laws?. *Journal of Finance*, *61*(1), 1-32.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. W. (1998). Law and finance. *Journal of Political Economy*, *106*(6), 1113-1155.

Lamont, O. A., & Thaler, R. H. (2003). Anomalies: The law of one price in financial markets. *The Journal of Economic Perspectives*, *17*(4), 191-202.

Lamont, O. A., & Thaler, R. H. (2003b). Can the market add and subtract? Mispricing in tech stock carve-outs. *Journal of Political Economy*, *111*(2), 227-268.

Lang, M. H., Lins, K. V., & Miller, D. P. (2003). ADRs, analysts, and accuracy: Does cross listing in the United States improve a firm's information environment and increase market value?. *Journal of Accounting Research*, *41*(2), 317-345.

Lee, C. M., Shleifer, A., & Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle. *Journal of Finance*, *46*(1), 75-109.

Lee, S. B., & Kim, K. J. (1993). Does the October 1987 crash strengthen the co-movements among national stock markets? *Review of Financial Economics*, *3*(1), 89.

Leung, A. C. M., Agarwal, A., Konana, P., & Kumar, A. (2016). Network analysis of search dynamics: The case of stock habitats. *Management Science*, *63*(8), 2667-2687.

Lewellen, J. (2002). Momentum and autocorrelation in stock returns. *Review of Financial Studies*, *15*(2), 533-564.

Li, E. X., Ramesh, K., Shen, M., & Wu, J. S. (2015). Do analyst stock recommendations piggyback on recent corporate news? An analysis of regular-hour and after-hours revisions. *Journal of Accounting Research*, *53*(4), 821-861.

Li, J., & Yu, J. (2012). Investor attention, psychological anchors, and stock return predictability. *Journal of Financial Economics*, *104*(2), 401-419.

Liu, H., & Peng, L. (2015). Investor attention: Seasonal patterns and endogenous allocations. *Working Paper.* Zicklin School of Business.

Liu, L. X., Sherman, A. E., & Zhang, Y. (2014). The long-run role of the media: Evidence from initial public offerings. *Management Science*, *60*(8), 1945-1964.

Lo, A. W., & MacKinlay, A. C. (1990). When are contrarian profits due to stock market overreaction?. *Review of Financial Studies*, *3*(2), 175-205.

Lou, D. (2014). Attracting investor attention through advertising. *Review of Financial Studies*, 27(6), 1797-1829.

Madsen, J. M., & Niessner, M. (2016). Is investor attention for sale? The role of advertising in financial markets. *Working Paper.* University of Minnesota and Yale School of Management.

Marosi, A., & Massoud, N. (2008). "You can enter but you cannot leave…": US securities markets and foreign firms. *Journal of Finance*, *63*(5), 2477-2506.

Mashruwala, C., Rajgopal, S., & Shevlin, T. (2006). Why is the accrual anomaly not arbitraged away? The role of idiosyncratic risk and transaction costs. *Journal of Accounting and Economics*, *42*(1-2), 3-33.

Menzly, L., & Ozbas, O. (2010). Market segmentation and cross-predictability of returns. *Journal of Finance*, *65*(4), 1555-1580.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance*, *42*(3), 483-510.

Mian, G. M., & Sankaraguruswamy, S. (2012). Investor sentiment and stock market response to earnings news. *Accounting Review*, *87*(4), 1357-1384.

Mitchell, M., Pulvino, T., & Stafford, E. (2002). Limited arbitrage in equity markets. *Journal of Finance*, *57*(2), 551-584.

Mondria, J. (2010). Portfolio choice, attention allocation, and price comovement. *Journal of Economic Theory*, *145*(5), 1837-1864.

Mondria, J., Wu, T., & Zhang, Y. (2010). The determinants of international investment and attention allocation: Using internet search query data. *Journal of International Economics, 82(1),* 85-95.

Morck, R., Yeung, B., & Yu, W. (2000). The information content of stock markets: why do emerging markets have synchronous stock price movements?. *Journal of Financial Economics*, *58*(1-2), 215-260.

Moskowitz, T. J., & Grinblatt, M. (1999). Do industries explain momentum?. *Journal of Finance*, *54*(4), 1249-1290.

Muslu, V., Rebello, M., & Xu, Y. (2014). Sell-side analyst research and stock Comovement. *Journal of Accounting Research*, *52*(4), 911-954.

Peng, L. (2005). Learning with information capacity constraints. *Journal of Financial and Quantitative Analysis*, *40*(02), 307-329.

Peng, L., & Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics*, *80(3),* 563-602.

Peress, J. (2014). The media and the diffusion of information in financial markets: Evidence from newspaper strikes. *Journal of Finance*, *69*(5), 2007-2043.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, *22*(1), 435-480.

Pindyck, R. S., & Rotemberg, J. J. (1993). The comovement of stock prices. *Quarterly Journal of Economics*, *108*(4), 1073-1104.

Pirinsky, C., & Wang, Q. (2006). Does corporate headquarters location matter for stock returns?. *Journal of Finance*, *61*(4), 1991-2015.

Pontiff, J. (1996). Costly arbitrage: Evidence from closed-end funds. *Quarterly Journal of Economics*, *111*(4), 1135-1151.

Pontiff, J. (2006). Costly arbitrage and the myth of idiosyncratic risk. *Journal of Accounting and Economics*, *42*(1-2), 35-52.

Pulatkonak, M., & Sofianos, G. (1999). The distribution of global trading in NYSE-listed non-US stocks. *Working Paper.* New York Stock Exchange.

Puthenpurackal, J. (2006). Explaining premiums in restricted DR markets and their implications: The Case of Infosys. *Financial Management*, *35*(2), 93-116.

Rabinovitch, R., Silva, A. C., & Susmel, R. (2003). Returns on ADRs and arbitrage in emerging markets. *Emerging Markets Review*, *4*(3), 225-247.

Rajgopal, S., & Venkatachalam, M. (2011). Financial reporting quality and idiosyncratic return volatility. *Journal of Accounting and Economics*, *51*(1-2), 1-20.

Ribeiro, R., & Veronesi, P. (2002). The excess comovement of international stock returns in bad times: a rational expectations equilibrium model. *Working Paper*. University of Chicago.

Roll, R. (1988). R-squared. *Journal of Finance*, *43*(2), 541-566.

Rosch, E., & Lloyd, B. B. (1978). Cognition and categorization, New Jersey: Lawrence Erlbaum.

Rosenthal, L., & Young, C. (1990). The seemingly anomalous price behavior of Royal Dutch/Shell and Unilever NV/PLC. *Journal of Financial Economics*, *26*(1), 123-141.

Ross, S. (1976). The arbitrage theory of capital asset pricing. Journal of Economic Theory, 13(3), 341-360.

Sarkissian, S., & Schill, M. J. (2003). The overseas listing decision: New evidence of proximity preference. *Review of Financial Studies*, *17*(3), 769-809.

Schmidt, D. (2013). Investors' attention and stock covariation: Evidence from Google sport searches. *Working Paper*. INSEAD.

Schultz, P., & Shive, S. (2008). Mispricing of dual class shares: Profit opportunities, arbitrage. and trading. *Working Paper*. University of Notre Dame.

Scruggs, J. T. (2007). Noise trader risk: Evidence from the Siamese twins. *Journal of Financial Markets*, *10*(1), 76-105.

Seasholes, M. S., & Wu, G. (2007). Predictable behaviour, profits, and attention. *Journal of Empirical Finance*, *14*(5), 590-610.

Shiller, R. (1989). Market volatility. Cambridge, Mass.: MIT Press.

Shiller, R. (2000). Irrational exuberance. Princeton, NJ: Princeton University Press.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*(3), 665-690.

Singh, P., Kumar, B., & Pandey, A. (2010). Price and volatility spillovers across North American, European and Asian stock markets. *International Review of Financial Analysis*, *19*(1), 55-64.

Skinner, D. J., & Sloan, R. G. (2002). Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. *Review of Accounting Studies*, *7*(2-3), 289-312.

Smith, B. F., & Amoako-Adu, B. (1995). Relative prices of dual class shares. *Journal of Financial and Quantitative Analysis*, *30*(2), 223-239.

Stambaugh, R. F., Yu, J., & Yuan, Y. (2015). Arbitrage asymmetry and the idiosyncratic volatility puzzle. *Journal of Finance*, *70*(5), 1903-1948.

Storms, K., Kapraun, J., & Rudolf, M. (2015). Can retail investor attention enhance market efficiency? Insights from search engine data. *Working Paper.* Otto Beisheim School of Management.

Suh, J. (2003). ADRs and US market sentiment. *Journal of Investing*, *12*(4), 87-95.

Teo, M., & Woo, S. J. (2004). Style effects in the cross-section of stock returns. *Journal of Financial Economics*, *74*(2), 367-398.

Teoh, S. H., Yang, Y. G., & Zhang, Y. (2009). R-square and market efficiency. *Working Paper.* University of California and The Chinese University of Hong Kong.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, *62*(3), 1139-1168.

Tetlock, P. C. (2010). Does public financial news resolve asymmetric information?. *Review of Financial Studies*, *23*(9), 3520-3557.

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance*, *63*(3), 1437-1467.

Veldkamp, L. L. (2006). Information markets and the comovement of asset prices. *Review of Economic Studies*, *73*(3), 823-845.

Vijh, A. M. (1994). S&P 500 trading strategies and stock betas. *Review of Financial Studies*, *7*(1), 215-251.

Vozlyublennaia, N. (2014). Investor attention, index performance, and return predictability. *Journal of Banking and Finance*, *41*, 17-35.

Wahal, S., & Yavuz, M. D. (2013). Style investing, comovement and return predictability. *Journal of Financial Economics*, *107*(1), 136-154.

Wilson, R. A., & Keil, F. C. (Eds.). (2001). *The MIT encyclopedia of the cognitive sciences*. MIT press.

Wurgler, J. (2000). Financial markets and the allocation of capital. *Journal of Financial Economics*, *58*(1-2), 187-214.

Yuan, Y. (2015). Market-wide attention, trading, and stock returns. *Journal of Financial Economics*, *116*(3), 548-564.

Zingales, L. (1995). Insider ownership and the decision to go public. *Review of Economic Studies*, *62*(3), 425-448.

Appendix

A2.1 Summary statistics for estimates from Equations (2.2a) and (2.2b)

$$V_{i,t}^{H} = a_{0i} + \beta_{1i}^{US} V_{i,t}^{US} + \beta_{2i} V_{i,t-1}^{H} + \beta_{3i} V_{i,t-1}^{US} + \beta_{4i,H} V_{H,t} + \beta_{5i,US} V_{US,t} + \beta_{6i,H} V_{H,t-1} + \beta_{7i,US} V_{US,t-1} + \varepsilon_{i,t} \qquad (2.2a)$$

$$V_{i,t}^{US} = a_{0i} + \beta_{1i}^{H} V_{i,t}^{H} + \beta_{2i} V_{i,t-1}^{US} + \beta_{3i} V_{i,t-1}^{H} + \beta_{4i,US} V_{US,t} + \beta_{5i,H} V_{H,t} + \beta_{6i,US} V_{US,t-1} + \beta_{7i,H} V_{H,t-1} + \varepsilon_{i,t} \qquad (2.2b)$$

	Panel A: Regression estimates from Equation (2.2a)								Panel B: Regression estimates from Equation (2.2b)							
	β_1^{US}	β_2	β_3	$\beta_{4,H}$	$\beta_{5,US}$	$\beta_{6,H}$	$\beta_{7,US}$	R^2	β_1^{H}	β_2	β_3	$\beta_{4,US}$	$\beta_{5,H}$	$\beta_{6,US}$	$\beta_{7,H}$	R^2
All																
Mean	0.3942	0.2186	-0.0080	0.6414	0.0555	-0.2117	0.0012	0.36	0.3893	0.2292	-0.0445	0.4483	0.0391	-0.1089	-0.0232	0.33
N	17,970	10,590	1,453	9,065	1,827	379	788		17,970	11,180	499	5,360	1,393	528	736	
Firm-qtrs	26,820	26,820	26,820	26,820	26,820	26,820	26,820		26,820	26,820	26,820	26,820	26,820	26,820	26,820	
Canada																
Mean	0.5459	0.1839	-0.0645	0.5984	0.0706	-0.1796	-0.0255	0.38	0.4267	0.2358	-0.0517	0.3567	0.1231	-0.0820	-0.0454	0.39
N	9,054	3,491	198	1,687	889	217	279		9,054	4,870	175	1,660	797	250	279	
Firm-qtrs	11,053	11,053	11,053	11,053	11,053	11,053	11,053		11,053	11,053	11,053	11,053	11,053	11,053	11,053	
Latin																
Mean	0.4314	0.1940	-0.0538	0.5342	0.0529	-0.0841	-0.0210	0.34	0.4162	0.2165	-0.0610	0.4756	0.0522	-0.0928	-0.0151	0.33
N	2,933	1,319	56	1,407	195	76	101		2,933	1,528	56	762	208	81	116	
Firm-qtrs	3,936	3,936	3,936	3,936	3,936	3,936	3,936		3,936	3,936	3,936	3,936	3,936	3,936	3,936	
Europe																
Mean	0.2562	0.2594	0.0221	0.6584	0.1338	-0.2336	0.0092	0.37	0.3927	0.2335	-0.0434	0.4416	-0.0713	-0.1120	0.0026	0.29
N	3,519	2,990	316	2,846	564	55	175		3,519	2,627	127	1,324	173	105	181	
Firm-qtrs	6,151	6,151	6,151	6,151	6,151	6,151	6,151		6,151	6,151	6,151	6,151	6,151	6,151	6,151	
Africa/Middle East																
Mean	0.2907	0.1920	0.1109	0.8113	-0.0603	-0.2202	0.0259	0.33	0.3197	0.2215	-0.0225	0.5720	0.0190	-0.1781	-0.0503	0.26
N	964	601	259	1,085	57	13	54		964	712	38	417	61	30	48	
Firm-qtrs	1,864	1,864	1,864	1,864	1,864	1,864	1,864		1,864	1,864	1,864	1,864	1,864	1,864	1,864	
Asia-Pacific																
Mean	0.1893	0.2918	0.0961	0.7663	-0.0548	-0.3970	0.0767	0.35	0.2818	0.2199	-0.0191	0.6354	-0.0299	-0.1643	0.0048	0.25
N	1,500	2,189	624	2,040	122	18	179		1,500	1,443	103	1,197	154	62	112	
Firm-qtrs	3,816	3,816	3,816	3,816	3,816	3,816	3,816		3,816	3,816	3,816	3,816	3,816	3,816	3,816	

A2.2 Definition of country-level and firm-level variables in Chapter 2

Attention Variables	
V^H	The volume shocks to the home-market shares of firm i on day t, and is computed based on the 200-day detrended measure on the stock's log-turnover as specified in Equation (2.1).
V^{US}	The volume shocks to the US cross-listed shares of firm i on day t, and is calculated similarly to V^H.
V_H	The aggregate volume shocks to the home-country market on day t, and is equal to the equally weighted volume shocks to all firms within that market with available data on a certain day (not including firm i), based on Thomson Reuters Datastream.
V_{US}	The aggregate volume shocks to the US market on day t, and is equal to the equally weighted volume shocks to all firms traded on NYSE, AMEX and NASDQ with available data on a certain day (not including firm i), based on CRSP data set .
AttentionComove	A proxy for investor attention comovement for firm i in quarter q, and is equal to the logarithmic transformation of the R^2, defined as $\ln(R^2/(1-R^2))$, obtained from the following model estimated for each firm-quarter: $V_{i,t}^H = a_{0i} + \beta_i^{US} V_{i,t}^{US} + \varepsilon_{i,t}$ which is equivalent to $V_{i,t}^{US} = a_{0i} + \beta_i^H V_{i,t}^H + \varepsilon_{i,t}$
Firm-level variables	
Earnings announcement speed	Following Gallemore and Labro (2015), measured as the number of days between the end of the fiscal year and the earnings announcement date, divided by 365 and multiplied by negative one.
Home number of analysts	The number of estimates in the home market underpinning the one-fiscal-year-ahead (FY1) earnings per share (EPS), as published in the IBES international file. The data is extracted at the monthly frequency and taken an average for each quarter.
US number of analysts	The number of estimates in the US market underpinning the one-fiscal-year-ahead (FY1) earnings per share (EPS), as published in the IBES US file. The data is extracted at the monthly frequency and taken an average for each quarter.
Home illiquidity	Calculated using Amihud (2002) illiquidity measure, which is the average ratio of the daily absolute return to dollar trading volume (measured in millions of dollars) in each quarter.
US illiquidity	Calculated similarly to Home illiquidity.
Institutional ownership	The percentage of a firm’s shares that are held by institutional investors, and is extracted from Thomson Reuters' 13F database at a quarterly frequency.
Dispersion of analysts	The standard deviation of earnings forecasts across all analysts divided by the absolute value of the mean estimate across all analysts. Due to the paucity of data available in the IBES US file, this measure is only computed for the home-market shares.
Return volatility	The standard deviation of daily returns of the home-market shares over the quarter.
ROA volatility	The standard deviation of return on assets (ROA) over the past 5 years with a minimum of 3 years observations. ROA is defined as the net income divided by average assets.
Market value	The natural logarithm of a firm's market capitalization in each quarter expressed in the US dollars.
Price	A firm's share price in the home market expressed in the US dollars.
Book-to-market ratio	A firm's balance sheet value of the common equity divided by market value of the common equity, and is extracted from Datastream at a quarterly frequency.

Idiosyncratic risk	Obtained by regressing, each quarter, daily return difference between the cross-listed pairs on the returns on the home-market index, returns on the US index, and relevant log currency changes (as specified in Equation (A2.1)). The standard deviation of the residual from the regression is defined as the idiosyncratic risk.
Dividend yield	The ratio of the total dollar value of the dividend paid over the quarter to the price of the share at the end of the quarter for the US cross-listed shares.
$Ln(P_{US}/P_H)$	The natural logarithm of the cross-listed share price expressed in dollars divided by the home-market share price also expressed in US dollars, and adjusted for the ADR bundling ratio.
$\lvert Ln(P_{US}/P_H)\rvert$	Calculated by averaging the daily price differences ($Ln(P_{US}/P_H)$)over a quarter, and taking the absolute value of the result.
Country-level variables	
Home-market volatility	Standard deviation of returns on home-country index, based on Thomson Reuters Datastream country index. Estimated quarter-by-quarter.
US-market volatility	Standard deviation of returns on Standard and Poor's 500 index. Estimated quarter-by-quarter.
Currency volatility	Standard deviation of foreign exchange rate for the home country, based on Thomson Reuters Datastream country exchange rate series. Estimated quarter-by-quarter.
Interest rate	A proxy for borrowing cost in the home market, measured as the annualized bank lending interest rate from the World Bank website.
Market integration	Correlation coefficient between returns on home-country index and returns on Standard & Poor's 500 index over the previous 60 months. Estimated month-by-month and taken an average for each quarter.
Financial development	The ratio stock market capitalization to GDP, extracted for each home country from the World Bank website at an annual frequency.
Stock market turnover	The ratio of the value of total shares traded to market capitalization, extracted for each home country from the World Bank website at an annual frequency.
Anti-director rights index	A country-level index compiled by La Porta et al. (1998), which measures the legal protection for minority shareholders in each market. The index is a sum of six antidirector rights scores, ranging from 0 to 5.
Disclosure index	A country-level index compiled by La Porta et al. (2006), which measures the disclosure requirements by the law or the listing rules of each financial market. The index is estimated from the arithmetic mean of six disclosure variables, including prospectus and disclosure on compensation, shareholders, inside ownership, contracts irregular, and transactions.

A2.3 Idiosyncratic risk

To measure idiosyncratic risk, we follow Gagnon and Karolyi (2010), for each firm-quarter, we estimate a time-series regression of the daily return difference on contemporaneous, leading, and lagged daily US- and home-market index returns, and the relevant log currency changes:

$$R_{US-H,t} = a + \sum_{i=-1}^{i=+1} \beta_i^{US} R_{M,t+i}^{US} + \sum_{i=-1}^{i=+1} \beta_i^{HM} R_{M,t+i}^{HM} + \sum_{i=-1}^{i=+1} \beta_i^{FX} R_{FX,t+i} + \varepsilon_{A-H,t} \qquad (A2.1)$$

where R_{A-H} is the return difference between the US cross-listed and the home-market shares ($\ln\left(\frac{P_t^{US}}{P_{t-1}^{US}}\right) - \ln\left(\frac{P_t^{H}}{P_{t-1}^{H}}\right)$) on day t.[64] $R_{M,t}^{HM}$ and $R_{M,t}^{US}$ are the US and home-market index returns, respectively. We use the local market index returns obtained from Datastream to proxy for home-maket returns, and S&P 500 index returns to proxy for US-market returns. R_{FX} is the returns on home-market currency relative to the US dollar calculated based on WM/Reuters WMR foreign exchange-rate quotes from Datastream. Idiosyncratic risk is the standard deviation of the residual from Equation (A2.1).

[64] Returns on both US- and home-market shares are calculated based on the US dollar share price. Home-market returns ($\ln\left(\frac{P_t^{H}}{P_{t-1}^{H}}\right)$) are calculated based on close-to-close prices. US-market returns ($\ln\left(\frac{P_t^{US}}{P_{t-1}^{US}}\right)$) are computed differently depending on the time-zone of the home market. When home-and US- markets are closed at the same time, US-market returns are calculated based on close-to-close prices. When home- and US- markets have imperfectly synchronized trading hours, US-market returns are calculated based on the midpoint of the bid/ask quotes at the time corresponding to the home-market closing time. When home- and US-markets have non-overlapping trading hours, US-market returns are based on the open-to-open quotes. A valid returns difference observation must meet two requirements: first, the stock has valid Datastream and TRTH prices on day t and t-1; second, the stock has non-zero trading volume on day t and t-1 in both US- and home-markets.

A3.1 Bellwether stocks in each industry

This table presents the details of bellwether stocks within each industry. Industry is defined using 48 Fama and French (1997) industry classifications. Bellwether firms are defined as those with a high analyst following and whose fundamentals are most reflective of other firms in the industry (see details in Section 3.4.3.1). For each industry and in each year, we count the number and calculate the average market capitalization of bellwether stocks and industry stocks, respectively. The table reports the average value over the sample period.

Industry	No. of bellwether stocks	Total No. of stocks	Average bellwether firm size ($billion)	Average firm size ($billion)
Food products	3	44	11.0211	3.8953
Recreation	1	24	7.5917	1.0885
Entertainment	3	47	10.6548	2.1338
Printing and publishing	1	16	1.9033	1.1063
Consumer goods	3	38	23.8343	8.8565
Apparel	2	33	3.9391	2.0797
Healthcare	5	68	4.5191	1.4053
Medical equipment	9	123	5.6723	1.6521
Pharmaceutical products	22	319	11.3039	4.2437
Chemicals	5	63	10.8635	4.4133
Rubber and plastic products	1	18	3.7136	1.3088
Construction materials	4	53	5.2234	1.7525
Construction	3	42	3.6386	1.4253
Steel works Etc	3	32	5.5767	1.6239
Machinery	8	100	10.6088	2.7306
Electrical equipment	3	56	19.5044	5.8397
Automobiles and trucks	4	44	8.0554	3.0796
Aircraft	2	20	26.4001	13.7031
Non-metallic and industrial	1	15	12.2027	4.3189
Coal	1	10	2.1816	1.8436
Petroleum and natural gas	11	151	33.7897	7.9638
Utilities	10	100	11.6123	6.0131
Communication	9	117	25.8206	8.0562
Personal services	3	38	1.9119	1.2347
Business services	30	405	10.7773	3.3380
Computers	8	110	6.5948	1.9825
Electronic equipment	15	211	17.8285	4.2018
Measuring and control	5	66	9.5886	2.7509
Business supplies	3	32	11.3958	4.9677
Shipping containers	1	11	4.7118	3.2458
Transportation	7	81	9.9866	3.6359
Wholesale	6	98	5.8818	1.9144
Retail	6	85	10.8433	3.4585
Restaurants, hotels, motels	4	58	11.2165	3.6194
Banking	35	571	9.1872	2.6982
Insurance	10	130	18.9336	6.9141
Real estate	1	25	4.6561	0.9160
Trading	8	119	14.0992	3.8918
Others	1	20	8.7807	2.9390

A3.2 Definition of key variables in Chapter 3

Variable	Definition
NewsTone	News tone for firm k, measured as the difference between positive and negative scores for each news and weighted by the news' relevance score to a firm.
NNEWS	The natural logarithm of one plus the total number of news articles covering the firm.
PCORR_ROA	Firm k's partial correlation in returns on asset (ROA) with all other firms in its industry, and is computed on a quarterly basis using quarterly ROAs over a five-year window as specified in Section 3.4.2.1.
PCORR_TONE	Firm k's partial correlation in news tone with all other firms in its industry, and is computed on a monthly basis using daily NewsTone over a three-month window as specified in Section 3.4.2.2.
PCORR_NNEWS	Firm k's partial correlation in news coverage with all other firms in its industry, and is computed similarly to PCORR_TONE.
LPCORR_ROA	The logarithmic transformation of PCORR_ROA, defined as: $Ln(PCORR\_ROA/(1-PCORR\_ROA))$.
LPCORR_TONE	The logarithmic transformation of PCORR_TONE.
LPCORR_NNEWS	The logarithmic transformation of PCORR_NNEWS.
RetComove	Firm k's return comovement in each month, and is estimated as the regression R^2 from the market model of returns using daily data over a three-month window, and then taken a log transformation as specified in Equation (3.13).
Analyst forecast revision	The earnings forecast revision of firm k in each month, calculated as the change in the mean forecast of 1-year ahead earnings per share from the previous month, scaled by firm k's stock price at the end of the previous month.
Forecast accuracy	The analyst forecast accuracy of firm k in each month, defined as the negative of the absolute value of the analyst forecast error (difference between actual earnings and median analyst forecast) deflated by the monthly stock price.
Volume shock	Trading volume shock of firm k in each month, measured as the difference between trading volume in the current month and the average trading volume over the past 12 months, scaled by the standard deviation of the trading volume over the past 12 months. Trading volume is defined as the volume trading for stock k in each month divided by the number of shares outstanding.
Market value	A firm's market capitalization in each month.
Book-to-market	A firm's book-to-market ratio calculated at the end of June each year following Fama and French (1992).
Price	A firm's share price in each month.
Turnover	The average daily share turnover over a month, where daily turnover is calculated as the number of trading volume on each day divided by the number of share outstanding.
Liquidity	Calculated using Amihud (2002) illiquidity measure, which is the average ratio of the daily absolute return to dollar trading volume (measured in millions of dollars) in each month.
Return volatility	The standard deviation of daily stock returns over a month.
Number of analysts	The number of estimates underpinning the one-fiscal-year-ahead (FY1) earnings per share (EPS), as published in the IBES.
Institutional ownership	The percentage of a firm's shares that are held by institutional investors, and is extracted from Thomson Reuters' 13F database at a quarterly frequency.

A4.1 Definition of key variables in Chapter 4

Variable	Definition
Analyst	The number of earnings forecast revisions made by sell-side analysts for a given firm, as published in the IBES.
News	The number of news articles issued by the business press for a firm, collected from TRNA.
AnalystShock	Firm *i*'s analyst shock in month *t*, measured as the number of analyst forecast revisions made for firm *i* in month *t* minus the average number of analyst forecast revisions over the past 12 months, as specified in Equation (4.1a).
NewsShock	Firm *i*'s news shock in month *t*, measured as the news coverage for firm *i* in month *t* minus the average news coverage over the past 12 months, as specified in Equation (4.1b).
Analyst R^2	The R^2 obtained from regressing firm *i*'s analyst forecast revisions on the equal-weighted style analyst forecast revisions (excluding firm *i*) using weekly data over a 12-month window.
News R^2	Calculated similarly to Analyst R^2, using News coverage as the attention measure.
Return R^2	The R^2 obtained from regressing the firm *i*'s weekly returns on the equal-weighted style returns in each week (excluding firm *i*) over a one-year window.
AnalystComove	The logarithmic transformation of Analyst R^2, defined as: $\ln\left(R^2/(1-R^2)\right)$
NewsComove	The logarithmic transformation of News R^2.
RetComove	The logarithmic transformation of Return R^2.
Size ($billions)	A firm's market capitalization at the end of June each year.
Price	A firm's share price at the end of June each year.
BM	A firm's book-to-market ratio calculated at the end of June each year following Fama and French (1992).
AbsRet	The absolute buy-and-hold returns over the period from July of year *t*-1 to June of year *t*.
Turnover	The average monthly turnover over the period from July of year *t*-1 to June of year *t*, where monthly turnover is the volume trading for stock *i* in each month divided by the number of share outstanding.
NUMEST	The number of estimates underpinning the one-fiscal-year-ahead (FY1) earnings per share (EPS), as published in the IBES.
Institutional ownership	The percentage of a firm's shares that are held by institutional investors, and is extracted from Thomson Reuters' 13F database at a quarterly frequency.
SalesGrowth	Sales growth measured as sales in year *t* divided by sales in year *t*-1.
ROA	Returns on asset measured as net income in year *t* scaled by total assets in year *t*-1.
StdROA	The standard deviation of annual return on assets between year *t*-4 and year *t*.
ROAComove	Comovement in ROA, measured as the log transformation of R^2 from a regression of firm *i*'s ROA on an equal-weighted style ROA (excluding firm *i*) using quarterly data over a 12-quarter window.