https://doi.org/10.1016/j.jbankfin.2019.105628

Aggregate investor sentiment and stock return synchronicity

This version: September 13, 2019

***Journal of Banking and Finance*, Forthcoming**

Timothy K. Chue*

School of Accounting and Finance

Hong Kong Polytechnic University

Kowloon, Hong Kong

Email: timothy.chue@polyu.edu.hk

Phone: (852) 2766-4995

Fax: (852) 2330-9845

Ferdinand A. Gul

School of Business

Deakin University

Melbourne, Australia

Email: ferdinand.gul@deakin.edu.au

G. Mujtaba Mian

College of Business

Zayed University, UAE

Email: Mujtaba.Mian@zu.ac.ae

*Corresponding author

Aggregate Investor Sentiment and Stock Return Synchronicity

Abstract

We show that the returns of individual stocks become more synchronous with the aggregate market during periods of high investor sentiment. We also document that the effect of sentiment on stock return synchronicity is especially pronounced for small, young, volatile, non-dividend-paying and low-priced stocks. This 'difference in difference' suggests that stocks with these characteristics are affected more by sentiment—consistent with previous studies. Our results support the hypothesis that greater constraints on arbitrage and the prevalence of sentiment-driven demand during periods of high sentiment lead to increased comovement among stocks.

JEL Classifications: G12; G14.

Keywords: Aggregate investor sentiment; Stock return synchronicity; Time-series variation; Cross-sectional difference.

1. Introduction

Stock return synchronicity, or the extent to which a stock co-moves with aggregate factors, is of fundamental concern to investors and portfolio managers. Not only does it directly affect the benefit of portfolio diversification, but the extent to which stocks move together also affects the potential payoff of stock selection. After all, the value of stock selection goes to zero if all stocks move in lockstep. As Engle (2002) emphasises, measures of comovement are critical inputs for many common tasks in financial management, from risk assessment to securities pricing. Brockman, Liebenberg and Schutte (2010) show that stock return synchronicity varies countercyclically over the business cycle, with changes in synchronicity accompanied by variations in information production. We complement their study by showing that even after these business cycle effects have been controlled for, aggregate investor sentiment still affects the variation of stock return synchronicity over time.

In addition to its portfolio implications, stock return synchronicity has been used as a proxy for the quantity of firm-specific information contained in stock prices. As proposed by Morck, Yeung and Yu (2000) and Durnev, Morck, Yeung and Zarowin (2003), individual stock returns become less synchronous with market returns when more firm-specific information is incorporated in stock prices.[1] Previous studies conduct *cross-sectional* (or cross-country) analyses that examine

[1] Studies in support of this interpretation include Wurgler (2000), Durnev, Morck and Yeung (2004), Piotroski and Roulstone (2004), Chan and Hameed (2006), Jin and Myers (2006), Chen, Goldstein and Jiang (2007), Ferreira and Laux (2007), Fernandes and Ferreira (2008, 2009), Brockman and Yan (2009), Francis, Huang, Khurana and Pereira (2009), Gul, Kim and Qiu (2010), Ferreira, Ferreira and Raposo (2011), Gul, Srinidhi and Ng (2011), Riedl and Serafeim (2011), Kim, Li and Li (2012), Xu, Chan, Jiang and Yi (2013) and Boubaker, Mansali and Rjiba (2014).

how stock return synchronicity is related to various factors that facilitate or inhibit informed trading, such as investor protection (Morck, Yeung, and Yu, 2000), managerial opportunism (Jin and Myers, 2006) and short-sale constraints (Bris, Goetzmann, and Zhu, 2007). We examine whether synchronicity changes over time and increases during periods when short-sale and other constraints to arbitrage are more binding (Shleifer and Vishny, 1997; Stambaugh, Yu, and Yuan, 2012). We investigate if there is *time-series* empirical support for Morck, Yeung and Yu's (2000) interpretation of stock return synchronicity. This investigation is important because a number of studies show that it is noise, rather than information, that drives idiosyncratic stock returns and question the use of stock return synchronicity as a measure of stock price informativeness.[2] Our finding that synchronicity increases during periods when investor sentiment plays a greater role in stock pricing thus contributes to this debate by presenting evidence consistent with the former interpretation of stock return synchronicity.

Baker and Wurgler (2006, 2007), Lemmon and Portniaguina (2006), Baker, Wurgler and Yuan (2012), and Huang, Jiang, Tu and Zhou (2015) show that aggregate investor sentiment forecasts a number of cross-sectional patterns in stock returns. Stambaugh, Yu and Yuan (2012) investigate the effect of sentiment on a broad set of cross-sectional anomalies in the stock market. Our work also belongs to this genre of research. Whilst previous literature focuses on the effect of sentiment on the *level* of aggregate prices and returns, we study whether aggregate sentiment also affects how these prices *co-move* with aggregate market factors – in a way that remains consistent with the idea that assets are less accurately priced during extreme sentiment periods. In this sense,

[2] Dasgupta, Gan and Gao (2010), Rajgopal and Venkatachalam (2011), and Bartram, Brown and Stulz (2012) find that firms that appear to be in a worse information environment have higher idiosyncratic return volatility.

our evidence contributes to the literature by providing independent, second-moment-based support for the hypothesis that aggregate investor sentiment affects asset prices at the aggregate level.[3]

To clarify the conceptual relationship between aggregate investor sentiment and stock return synchronicity, we use a modified version of the model developed by Daniel, Hirshleifer and Subrahmanyam (1998). As in Daniel, Hirshleifer and Subrahmanyam (1998), investors can become overconfident in the precision of the signals they receive. Here, we further assume that sentiment affects the extent to which investors overestimate signal precision. In particular, when *aggregate* investor sentiment is strong, investors overestimate the precision of the *market* signal—leading to over-reaction to the signal and "excess co-movement" across stocks (relative to a scenario in which sentiment is normal and no overreaction occurs). We present this model in Appendix B.

Empirically, we examine the association between investor sentiment and stock return synchronicity using the investor sentiment index constructed by Baker and Wurgler (2006, 2007). Although Baker and Wurgler already orthogonalise their index with respect to several macroeconomic variables, Sibley, Wang, Xing and Zhang (2016) argue that the resulting index is still correlated with business cycle variables. They show that about 41% of the variation in the Baker-Wurgler index can be explained by the T-bill rate and Lee's (2011) liquidity risk factor. To

[3] There are studies that examine the effect of investor sentiment on other aspects of the aggregate stock market. Yu and Yuan (2011) and Antoniou, Doukas and Subrahmanyam (2015) examine, respectively, how aggregate investor sentiment influences the market's mean-variance trade-off and the validity of the capital asset pricing model. Antoniou, Doukas and Subrahmanyam (2013) focus on how sentiment affects the returns on the momentum strategy. In relation to economic policy, Kurov (2010) shows that investor sentiment plays a significant role in the effect of monetary policy on the stock market.

address the concern that the Baker-Wurgler index captures variations in macroeconomic conditions, we follow the procedure discussed in Sibley, Wang, Xing and Zhang (2016) and use a residual sentiment index that has been orthogonalised to these variables for our analyses. We follow Morck, Yeung and Yu (2000) and compute our measure of stock return synchronicity based on a logistic transformation of the R-squared in regressions with individual stock returns as the dependent variable. However, as opposed to simply using market and industry returns as explanatory variables in the regressions, we use the Fama-French (1993) three factors and momentum instead. Ang, Hodrick, Xing and Zhang (2006) point out that the residuals from such a model serve as better measures of idiosyncratic volatility. As in Morck, Yeung and Yu (2000), we aggregate the stock-level measure of synchronicity to obtain a market-wide measure of stock return synchronicity.

In our base case analysis, we regress aggregate market-wide stock return synchronicity on sentiment and other time-series variables. Motivated by the findings of Yu and Yuan (2011), Stambaugh, Yu and Yuan (2012) and Antoniou, Doukas and Subrahmanyam (2013, 2015), we allow the effect of sentiment on synchronicity to be asymmetric. Due to short-sale constraints, overpricing during high sentiment periods is more prevalent than underpricing during low sentiment periods. Indeed, we find that stock return synchronicity significantly increases when positive investor sentiment becomes more bullish but not when negative sentiment turns more bearish. We also find that *both* GDP growth rate and investor sentiment remain significantly related to the variation in return synchronicity, even when both variables are included in the multiple regressions. As stock return synchronicity *declines* during business cycle expansions but *increases* in bullish sentiment states, our results cannot simply arise from the misclassification of business cycle booms as states of bullish investor sentiment. Our evidence therefore uncovers a behavioural source of time-varying stock return synchronicity that is distinct from the rational source examined by Veldkamp (2005, 2006) and Brockman et al. (2010), who focus on endogenous information signals generated from aggregate economic activities.[4]

To show that information plays a role in the link between aggregate sentiment and stock return synchronicity, we also use the private-information-based trading (*PrInfo*) measure proposed by Llorente, Michaely, Saar and Wang (2002) to proxy for the extent to which informed trading affects stock prices.[5] We calculate *PrInfo* over different periods and find that, indeed, stock prices are driven more by behavioural trading during high sentiment periods.

As further support that time-series variations in stock return synchronicity are driven by investor sentiment, we turn to cross-sectional evidence. We find that the effect of sentiment on stock return synchronicity exhibits a cross-sectional pattern, disproportionately affecting companies with certain characteristics. As Baker and Wurgler (2006, 2007) and Lemmon and Portniaguina (2006) emphasise, such a cross-sectional pattern is key to demonstrating that it is really investor sentiment and limits to arbitrage that are at work. In particular, companies that are more difficult to arbitrage and whose valuations are more subjective (such as those that are small, young, volatile, non-dividend paying, low-priced and with extreme valuation ratios, as suggested

[4] Bekaert and Hoerova (2016) also show that the Baker-Wurgler sentiment index does not capture risk aversion changes.

[5] This measure relies on return continuation versus reversals following unusual volume days as proxies for the degree of informed trading in the stock market. Llorente et al. (2002) argue that if the unusual volume on a given day is followed by return continuation over subsequent days, the unusual volume is likely driven by information. In contrast, if the unusual trading activity is followed by return reversals, the unusual trading is likely due to non-informational reasons.

by Baker and Wurgler, 2006) tend to be more exposed to sentiment-based demand. We find that these firms indeed experience more substantial increases in stock return synchronicity following periods of high investor sentiment. We see this difference in difference by first sorting stocks according to their characteristics and then calculating their return synchronicity for different sentiment periods. We also carry out pooled, time-series and cross-sectional regressions to show that this conclusion remains in a multivariate setting in which GDP growth rate, return on equity (ROE) synchronicity and different firm-level measures of stock liquidity are controlled for. Further evidence for this cross-sectional variation comes from the fact that our base case results for the aggregated market-wide synchronicity become insignificant when we value-weight firm synchronicity in computing aggregated market-level synchronicity, indicating that sentiment has little effect on the returns of large, mature companies that are easier to value and face fewer arbitrage constraints.

What can explain the observed relationship between investor sentiment and stock return synchronicity? The financial press suggests such a link. 'Sentiment greatly reduces the value of picking individual stocks, as low-quality names often move along with higher value shares. … In such an environment, … investors care less about distinguishing the differences between individual companies'.[6] This quote summarises the view that during periods of extreme investor sentiment in the stock market, investors become more prone to the prevailing sentiment, and their desire to distinguish between winners and losers declines. With less firm-specific information reflected in prices, return synchronicity increases.

[6] 'Time to abandon herd mentality', *Wall Street Journal*, September 15, 2009.

Interest from the popular press aside, our study of the link between investor sentiment and stock return synchronicity is also motivated by the relevant academic literature. Theoretically, studies on the limits of arbitrage show that informed arbitrage tends to become more constrained and stock prices tend to become less informative during periods of extreme investor sentiment. For example, Shleifer and Vishny (1997) model the agency problems associated with delegated portfolio management and show that arbitrageurs curtail their information-based trading during extreme sentiment periods, when prices are farthest from fundamental values. A number of studies show that sentiment-driven demand can cause investors to become less discriminating among firms of different quality (Cooper, Dimitrov, and Rau, 2001; Cooper, Khorana, Osobov, Patel, and Rau, 2005; Greenwood and Hanson, 2013). Lastly, the idea that stock return synchronicity increases when stock prices reflect less firm-specific information is in line with the evidence presented by Morck, Yeung and Yu (2000), Durnev, Morck, Yeung and Zarowin (2003), Durnev, Morck and Yeung (2004) and Chen, Goldstein and Jiang (2007), among others.

The rest of this paper proceeds as follows. In Section 2, we describe our data and discuss the construction of the variables to be used in the subsequent analyses. Section 3 reports our empirical results. We first provide evidence of the relationship between investor sentiment and stock return synchronicity aggregated at the market level, followed by how this relation varies in the cross-section. Section 4 reports a number of robustness checks on our main findings. Section 5 concludes the paper.

2. Data description and variable construction

This paper examines the relationship between stock return synchronicity and investor sentiment using both time-series and panel regression frameworks. Our primary data consist of common stocks traded on the NYSE, AMEX and NASDAQ stock exchanges. Information on these common stocks, which have share codes 10 and 11, is obtained from the Center for Research in Stock Prices (CRSP). We obtain quarterly financial statement data for these stocks from Compustat. The Baker-Wurgler index of aggregate investor sentiment is provided by Jeffrey Wurgler.[7] In robustness checks, we also examine the partial least squares (PLS) based investor sentiment index constructed by Huang, Jiang, Tu and Zhou (2015).[8]

Our basic empirical approach involves measuring stock return synchronicity each month, at either the individual stock or aggregate market level, and relating it with lagged values of monthly investor sentiment. Our period of analysis ranges from February 1966 to January 2011, a total of 540 months. The remainder of this section provides further details on variable construction.

2.1. Measures of stock return synchronicity

We use the *R*-squared of Carhart's (1997) four-factor model (which includes the Fama-French (1993) three factors and the momentum factor) to construct estimates of return synchronicity at both the individual firm and market levels. Although Morck, Yeung and Yu (2000, MYY henceforth) use a simpler market model to construct measures of synchronicity, Ang,

---

[7] We thank Jeffrey Wurgler for making available the data on his web site (http://pages.stern.nyu.edu/~jwurgler/).

[8] We thank Guofu Zhou for making the PLS-based investor sentiment index available on his web site (http://apps.olin.wustl.edu/faculty/zhou/).

Hodrick, Xing and Zhang (2006) point out that the residuals from the Fama-French factors better measure idiosyncratic volatility. The factor model also allows individual stocks to have different betas with respect to the factors, which can potentially be an important source of variations in stock return synchronicity. This is an attractive feature in our context because we are especially interested in how the relation between sentiment and return synchronicity varies across stocks with different characteristics (such as size and age), which can have different betas. Our results remain robust when we use the MYY model (which includes market and industry factors), the Fama and French (2015) five-factor model or a global four-factor model (which includes global market, size, value and momentum factors) to compute stock return synchronicity.[9]

For each stock in each month, we regress daily stock returns on the Fama-French three factors and momentum:

$$r_{j.d} = \beta_0 + \beta_{mkt,j}\, MKT_d + \beta_{HML,j}\, HML_d + \beta_{SMB,j}\, SMB_d + \beta_{UMD,j}\, UMD_d + \varepsilon_{j,d}, \tag{1}$$

where $r_{j,d}$ is the return of stock j on day d, and the explanatory variables are the standard Fama-French three factors – market (*MKT*), value (*HML*) and size (*SMB*) – together with the momentum factor (*UMD*). To have a sufficient number of daily observations in the estimation of monthly return synchronicity, we remove stock-month observations in which a stock has missing returns in more than 10% of that month's trading days.

[9] All of the data on factor returns are downloaded from Kenneth French's Data Library at the website http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. We thank Kenneth French for making these data available.

The coefficient of determination, R^2, of Equation (1) is a measure of synchronicity between stock j's return and the factor returns. This measure, however, is bounded between 0 and 1 and suffers from the econometric issues associated with large skewness and kurtosis. To circumvent these issues, we apply a logistic transformation to raw R^2 to obtain our measure of return synchronicity, $Synchronicity_{j,t}$, for stock j in month t.

$$Synchronicity_{j,t} = \ln\left(\frac{R_{j,t}^2}{1-R_{jt}^2}\right) \tag{2}$$

To convert this stock-specific measure into an aggregate measure of return synchronicity for the market, or for a subgroup of stocks, we follow Morck, Yeung, and Yu (2000) and first compute an aggregated R^2 for month t:

$$R_t^2 = \frac{\sum_j R_j^2 \times SST_j}{\sum_j SST_j}, \tag{3}$$

where SST is the sum of squared total variations for stock j obtained from Equation (1).[10] The logistic transformation of the expression in Equation (3) yields an aggregate measure of return synchronicity, $Synchronicity_t$, for month t.

$$Synchronicity_t = \ln\left(\frac{R_t^2}{1-R_t^2}\right) \tag{4}$$

[10] We note that this commonly used weighting scheme assigns greater weight to volatile stocks than stable stocks. In a later section, we discuss results that use market capitalizations as weights.

The summary statistics for *Synchronicity* are reported in Table 1, Panel A. The distribution for stock-month observations has been winsorised at the 1st and 99th percentiles to mitigate the effect of outliers on our analyses.

2.2. Measure of investor sentiment

We use the Baker-Wurgler (2006, 2007) sentiment index as our primary measure of investor sentiment. The index is based on six measures of investor sentiment: closed-end fund discount, NYSE share turnover, number of IPOs, first day returns on IPOs, share of equity issues in total debt and equity issues, and dividend premium (the log difference of the average market-to-book ratios of payers and non-payers). Baker and Wurgler orthogonalise each of these measures with respect to a number of macroeconomic and business cycle variables and then extract the first principal component as the sentiment index. The index is available monthly.

Despite Baker and Wurgler's (2006) orthogonalisation procedure, Sibley, Wang, Xing, and Zhang (2016) point out that the resulting index still contains information about economic fundamentals that needs to be purged before it can truly capture investors' behavioural biases. They show that 63% of the variation in the Baker-Wurgler index can be explained by 13 contemporaneous economic and financial market variables. To alleviate concerns of over-fitting, Sibley et al. (2016) show that just two variables, the T-bill rate and Lee's (2011) liquidity factor, can capture the bulk of this correlation – about 41% of the variation in the Baker-Wurgler index can be explained by these two variables.[11] Sibley et al. (2016) thus suggest that the Baker-Wurgler index be orthogonalised with respect to either all 13 variables or, to avoid concerns of over-fitting, with respect to the T-bill rate and the Lee (2011) liquidity risk factor.

We follow the procedure in Sibley, Wang, Xing and Zhang (2016) closely and regress the Baker-Wurgler index on the T-bill rate and Lee's (2011) liquidity risk factor (2011). This procedure allows us to decompose the Baker-Wurgler index into a component that is related to the T-bill rate and the Lee (2011) liquidity risk factor and a residual component that is orthogonal and may better capture behavioural biases. We use this residual component as our primary measure of sentiment throughout this paper. Our results remain robust if we use the original Baker-Wurgler index instead.

Both the Baker-Wurgler index (denoted *Sent*) and the residual sentiment measure (denoted *SentRes*) have means of zero by construction. Values close to zero thus indicate sentiment that is neither bullish nor bearish. In contrast, values that are significantly away from zero indicate extreme sentiment: large positive values indicate bullish sentiment and large negative values represent bearish sentiment. We report their summary statistics in Table 1, Panel B. The means of these variables do not equal zero as our sample start date of January 1966 is subsequent to the July 1965 start date for the Baker-Wurgler index. To examine the link between stock return synchronicity and investor sentiment, we match return synchronicity in month t with the sentiment index in month t-1.

[11] Sibley et al. (2016) acknowledge that the critical assumption underpinning their analysis is that the variables they identify primarily represent macroeconomic fundamentals and are not influenced by investor sentiment.

The correlation matrix in Table 2 shows that the original Baker-Wurgler sentiment index, *Sent*, has a correlation of 0.796 with the residual index (*SentRes*). The positive correlation between *Sent* and *Tbill* (the T-bill rate) and the negative correlation between *Sent* and *PctZero* (a measure of market illiquidity based on zero-return days) are both highly significant. However, by construction, the correlations between *SentRes* and these variables are no longer statistically significant.

2.3. Control and cross-sectional rank variables

Even though the residual Baker-Wurgler (2006, 2007) index we use has been orthogonalised with respect to four major macroeconomic factors as well as the T-bill and market-wide liquidity factor of Lee (2011), we add two control variables in our regression analyses to account for fundamentals-driven time variations in stock return synchronicity. First, following Brockman, Liebenberg and Schutte (2010), we include the lagged real GDP growth rate in all our specifications. Brockman et al. (2010) show that stock return synchronicity is countercyclical: stronger during recessions and weaker in expansions. Second, we construct a measure of ROE synchronicity for each firm-quarter, using data from the most recent 12 quarters. We regress the ROE of a firm on contemporaneous market and industry ROE. The logistic transformation of the R^2 on this regression, analogous to Equation (2), yields our ROE synchronicity measure. We provide more details on the construction of this variable in Appendix A. As these controls are only available at quarterly frequencies, we use the values from their most recently available quarterly observations in our monthly regressions. We also aggregate this firm-level measure of *ROE Synchronicity* into a market-wide *ROE Synchronicity* measure, while ensuring that the weight assigned to each stock is the same as that in the computation of aggregate return synchronicity in Equation (3). Specifically, we first aggregate the R^2 of the individual firm ROE regressions into a market-wide R^2 using Equation (3), with *SST* coming from the corresponding return regressions. We then use Equation (4) to compute market-wide *ROE Synchronicity* for a given quarter. Table 1, Panels C and D, respectively, report summary statistics of the market-wide and firm-level control variables.

Table 2 shows that all measures of market-level return synchronicity are negatively associated with aggregate stock market illiquidity – as measured by *PctZero* – consistent with the cross-sectional findings of Chan, Hameed and Kang (2013). Table 2 also shows that the different synchronicity measures are negatively related to the GDP growth rate, consistent with the results reported by Brockman, Liebenberg and Schutte (2010).

The literature on investor sentiment argues that the effect of sentiment is not uniform across stocks and varies predictably in a cross-section of stocks. Stocks that have more subjective valuations and are more difficult to arbitrage are disproportionately affected by sentiment. To identify such stocks, the literature relies on a number of firm characteristics. Firm size is perhaps the most commonly used variable in this regard – it is commonly argued that small stocks are more affected by sentiment than large stocks (Lemmon and Portniaguina, 2006; Qiu and Welch, 2006; and Baker and Wurgler, 2006). In addition, Baker and Wurgler (2006) document that investor sentiment has a much greater effect on the prices of young, volatile, non-dividend paying and extreme market-to-book ratio stocks than the prices of mature, stable, high-dividend paying and medium market-to-book stocks. Finally, the results in Brandt, Brav, Graham and Kumar (2010) suggest that the stock return synchronicity of low-priced stocks may be affected more by sentiment because trading in such stocks is dominated by small investors.

To examine whether the effect of sentiment on stock return synchronicity varies in a cross-section, we sort stocks into subgroups at the end of each month according to the six characteristics noted above: size, age, stock return volatility, dividend-to-book ratio, market-to-book (MB) ratio and stock price level. A detailed description of how we compute these characteristics and how we sort stocks into subgroups is outlined in Appendix A, Panel D. In brief, we sort stocks using breakpoints determined from the NYSE stock universe, and ranks are assigned in a way such that low-ranked stocks are those that are relatively more difficult to value and arbitrage. A simple sorting exercise for size, age, dividend payout and stock price ensures that lower ranks are assigned to stocks that are more prone to sentiment. We reverse the sort order for stock return volatility to ensure the same holds for stocks sorted on return volatility. When sorting stocks on dividend-to-book ratio, we first assign stocks that do not pay any dividend to decile 0. The remaining dividend-paying stocks are then grouped into deciles 1 through 10 based on their dividend-to-book ratio.

Finally, the MB ratio differs from the other characteristics in that, as Baker and Wurgler (2006) argue, both extremely high and extremely low MB stocks can be more exposed to sentiment. We, therefore, first sort stocks into deciles based on their MB ratio and then re-assign them into ranks such that low-ranked stocks are expected to be more prone to sentiment than high-ranked stocks. We assign a rank of 1.5 to stocks in the top and bottom deciles, a rank of 3.5 to stocks in deciles 2 and 9, and so on. Stocks in the middle two deciles are assigned a rank of 9.5. This procedure ensures that all of our ranks can be interpreted consistently.

Rather than separately reporting results for each of the characteristics, which are undoubtedly correlated with each other, to conserve space, we construct a composite measure based on the individual characteristics. We then examine how the effect of sentiment on synchronicity varies with this composite measure. Following Mian and Sankaraguruswamy

(2012), the composite measure is based on the first principle component of the decile ranks of the individual characteristics. We sort stocks each month into decile ranks based on this composite measure, which is denoted as *Rank6V*. Stocks in the bottom (top) decile are the most (least) exposed to the effects of sentiment. However, it is worth noting that our cross-sectional results hold individually for each of the six characteristics.

Because cross-sectional variations in stock liquidity can influence cross-sectional variations in stock return synchronicity (Chan, Hameed, and Kang, 2013), we also compute two stock-level measures of illiquidity and include them alternately as controls in all our cross-sectional analyses. The first is the percentage of zero-return days (*PctZero*) in a given month, as introduced by Lesmond, Ogden and Trzcinka (1999). The second is the average daily Amihud (2002) illiquidity over a month, in which daily illiquidity of a stock is computed as the stock's absolute daily return divided by its daily dollar volume. We compute these measures of illiquidity for each stock in each month. Higher values of these measures indicate greater illiquidity.

3. Empirical results

3.1. Relation between investor sentiment and aggregate stock return synchronicity

To investigate the relation between the aggregate market-level measure of stock return synchronicity, computed in Equation (4), and the residual investor sentiment index, we run regressions of the following form.

$$Synchronicity_t = a_0 + a_1 SentResPos_{t-1} + a_2 AbsSentResNeg_{t-1} + a_3 GDP_{t-1} + a_4 PeriodCount_t + a_5 PeriodCountSq_t + a_6 ROESynchronicity_{t-1} + \varepsilon_t \tag{5}$$

The key explanatory variables are *SentResPos* and *AbsSentResNeg*, which allow us to capture the asymmetric effect of investor sentiment on synchronicity. *SentResPos* takes the value of the residual sentiment index, *SentRes*, when it is positive and zero otherwise. Analogously, *AbsSentResNeg* takes the absolute value of *SentRes* when it is negative and zero otherwise. As values of sentiment close to zero represent sentiment that is neither particularly bullish nor bearish, large positive and negative values indicate periods when sentiment is high (i.e. bullish) and low (i.e. bearish). Taking the absolute value of sentiment in defining *AbsSentResNeg* makes the expected coefficients on both *SentResPos* and *AbsSentResNeg* positive. If the effect of sentiment on synchronicity is symmetric, moving from 'normal' to both extremely high and extremely low sentiment leads stock return synchronicity to increase. However, Yu and Yuan (2011), Stambaugh, Yu and Yuan (2012) and Antoniou, Doukas, and Subrahmanyam (2013, 2015) present evidence of an asymmetric effect of sentiment, whereby stocks become mispriced primarily during periods of high sentiment but not during periods of low sentiment. If similar asymmetry is to be observed in the effect of sentiment on synchronicity, we expect *SentResPos*, but not *AbsSentResNeg*, to be significantly positive.

GDP denotes the growth rate of real GDP. The inclusion of this variable is to account for the time variation in stock return synchronicity that is related to general economic activities, as documented by Brockman, Liebenberg and Schutte (2010). The other time-series control variables are *PeriodCount* and *PeriodCountSq*. *PeriodCount* is a linear time trend variable that takes the value of one in the first month of our sample (i.e. October 1966) and increases by one in each subsequent month. In the last month of our sample (i.e. January 2011), it takes the value of 532. *PeriodCountSq* is simply the square of *PeriodCount*. Together, these two variables aim to control for the long-run trend in idiosyncratic return volatility (as documented by Campbell, Lettau, Malkiel, and Xu, 2001) and its reversal (as documented by Brandt, Brav, Graham, and Kumar, 2010). It is worth noting that a follow-up study by Bekaert, Hodrick and Zhang (2012) shows that the volatility trend in the US market is not permanent and is likely the result of temporary switching of volatility to a higher variance regime.[12] Nevertheless, to be conservative, we include these controls in our analysis. We control for fundamentals-driven stock return synchronicity by including the variable *ROE Synchronicity*, which measures the synchronicity between firm-level ROE and market- and industry-level ROE. In a simple rational model, we expect the synchronicity of stock returns to be associated with the synchronicity of firm fundamentals. Empirical relevance of such a link is underscored by studies such as Wei and Zhang (2006), Irvine and Pontiff (2009), and Bekaert, Hodrick and Zhang (2012), among others.[13] Finally, it is worth reiterating that we already control for the effect of market-wide liquidity on synchronicity by employing the residual measure of Baker-Wurgler sentiment that has been orthogonalised to Lee's (2011) measure of market-wide illiquidity.

---

[12] They also show that there is no significant trend in volatility in non-US, developed equity markets.

[13] The primary purpose of many of these studies is to rationalize the earlier documented evidence of time trend in idiosyncratic volatility in the US stock market. They differ in terms of the variables they choose to operationalize the notion of fundamentals-driven volatility. We use *ROE* to account for time-varying synchronicity of firm fundamentals, as in Wei and Zhang (2006). We do not include additional variables linked to the time trend, as we directly add the two trend variables, *PeriodCount* and *PeriodCountSq*, as controls.

Table 3, Panel A reports the estimated coefficients of this regression, along with *t*-statistics based on Newey-West (1987) standard errors. We begin by examining the relationship between the MYY measure of return synchronicity (i.e. with market and industry returns as factors) and the Baker-Wurgler index, *Sent.* To allow for potential asymmetric effects of positive and negative sentiment, we use *SentPos* and *AbsSentNeg*, which take the positive and negative absolute values of *Sent*, respectively. Table 3, Panel A, Column (1) reports these results. Our results are supportive of an asymmetric relationship between investor sentiment and return comovement as only *SentPos* (not *AbsSentNeg*) is significantly positive. Consistent with the findings of Brockman, Liebenberg and Schutte (2010), we also find that higher GDP growth rates are associated with lower stock return comovement.

We then turn to the use of the Fama-French three factors plus momentum to compute stock return synchronicity. These results are reported in Columns (2) and (3). Column (2) still uses *Sent*, the basic Baker-Wurgler index, to measure investor sentiment. Column (3) uses *SentRes*, which has been orthogonalised with respect to the T-bill rate and the Lee (2011) liquidity factor. Regardless of whether *Sent* or *SentRes* is used, we see that investor sentiment has a significant effect on stock return synchronicity. We also see that the effect is asymmetric, only *SentPos* and *SentResPos* (not *AbsSentNeg* and *AbsSentResNeg*) are significantly positive.

As we discuss above, the Fama-French and momentum factors can better capture market-wide comovement than the MYY factors, and *SentRes* may be a cleaner measure of investor sentiment than *Sent*. We thus use the specification in Column (3) as our baseline specification. To gauge the economic significance of our baseline findings, we compare the effects of *SentRes* with those of *GDP* on stock return synchronicity. From Table 1, we see that the standard deviation of *SentRes* and *GDP* are comparable, yet the magnitude of the coefficient on *SentResPos* is more than three times that on *GDP*, suggesting that *SentRes* is an even more significant determinant of stock return synchronicity than *GDP*.

We next examine whether the relation between sentiment and stock return synchronicity is sensitive to the use of alternative factor models to estimate stock return synchronicity. The inclusion of additional factors could be important if investor sentiment were to affect certain omitted systemic factors. To examine this possibility, we repeat our analysis and compute stock return synchronicity using the Fama-French (2015) five-factor model, which includes the Fama-French (1993) three factors (market, size and book-to-market) and profitability and investment factors. Table 3, Panel A, Column (4) shows that, even when the Fama-French five factors are used, the coefficient on *SentResPos* continues to be positive and significant and the coefficient on *AbsSentResNeg* continues to be insignificant, suggesting that the asymmetric relation between sentiment and synchronicity remains robust to this alternative measure of stock return synchronicity.

Another potential concern is that the factor models we use ignore global factors, which may be especially important during more recent time periods when international markets have become better integrated. To alleviate this concern, we compute stock return synchronicity from a global four-factor model, with global versions of the Fama-French three factors and the momentum factor. As these data are only available starting from 1991, we use the global four factors to estimate return synchronicity from 1991 onwards and the local (US) four factors for the pre-1991 period. This approach seems reasonable as a long time series is needed to investigate the effects of sentiment over time, and the effects of global factors on the US market should be less pronounced in earlier years. In Table 3, Panel A, Column (5), we report results using this alternative measure of stock return synchronicity. The coefficients on *SentResPos* and

AbsSentResNeg remain largely unaffected, suggesting that omitted global factors are unlikely to affect the relation between aggregate investor sentiment and stock return synchronicity that we document.

A number of recent studies show that *style-specific* sentiment increases the comovement of returns among stocks that belong to the same style (Barberis and Shleifer, 2003; Barberis, Shleifer, and Wurgler, 2005). In contrast, the investor sentiment that we analyse, instead of being specific to investors within a particular style or 'habitat', is measured at the *aggregate* level. As emphasised by Baker and Wurgler (2007, p. 131-2), such a sentiment measure can be interpreted as investors' 'propensity to speculate' or simply their 'optimism or pessimism about stocks in general'. While previous studies show, *cross-sectionally*, groups of stocks favoured by the same clientele tend to move together, we find that, *over time*, return synchronicity with the aggregate stock market strengthens when the intensity of aggregate investor sentiment is high.

3.2. Relation between investor sentiment and the prevalence of informed trading

To further delve into whether the link between sentiment and stock return synchronicity is related to the information content of stock prices, we use the private-information-based trading (*PrInfo*) measure advocated by Llorente, Michaely, Saar, and Wang (2002). This measure relies on return continuation versus reversals following unusual volume days. Llorente et al. (2002) argue that if unusual volume on a given day is followed by return continuation over subsequent days, then the unusual volume is likely driven by information. In contrast, if the unusual trading activity is followed by return reversals, then it is likely caused by non-informational reasons.[14] Specifically, *PrInfo* is given by b_2 in the following regression,

$$r_{j,d} = b_0 + b_1 r_{j,d-1} + b_2 r_{j,d-1} V_{j,d-1} + \varepsilon_{j,d} \tag{6}$$

where $r_{j,d}$ is the return for stock j on day d, and $V_{j,d-1}$ is the day d-1 log turnover for stock j, detrended by subtracting a 200-trading-day moving average. We run the regression given by Equation (6) above to obtain an estimate of *PrInfo* for each stock in each month, using daily return and volume data. A positive (negative) value for *PrInfo* indicates that the daily abnormal volume for a stock during the month is, on average, associated with return continuation (reversals), implying the prevalence of informed (behavioural) trading. We winsorise the pooled, firm-month distribution of *PrInfo* at the 1st and 99th percentiles and report its summary statistics in Table 4, Panel A, in which all numbers have been multiplied by 100 for ease of exposition. The mean (median) value of *PrInfo* is -2.20 (-0.54), which is comparable to the statistics reported by Llorente et al. (2002) in their Tables 2 and 3. The standard deviation of 37.30 is rather large, which reflects cross-sectional variations linked to differences in factors such as size and market microstructure (Llorente et al., 2002). To arrive at an aggregate measure of *PrInfo* for a month, we take a simple average of *PrInfo* across all stocks in that month. As expected, this aggregated monthly time series

[14] The behavioural finance literature in the past two decades has included significant developments that help understand these ‘non-informational reasons’ for trading. Overconfidence, over-extrapolation, and prospect-based preferences have been established as key cognitive biases that drive the behaviour of irrational traders in financial markets. See, for example, Barberis (2018) for an updated discussion.

of *PrInfo* exhibits much less volatility, with a standard deviation of 3.22. We then estimate Equation (5) by replacing stock return synchronicity with *PrInfo*:

$$PrInfo_t = b_0 + b_1 SentResPos_{t-1} + b_2 AbsSentResNeg_{t-1} + b_3 GDP_{t-1} + b_4 PeriodCount_t + b_5 PeriodCountSq_t + b_6 ROESynchronicity_{t-1} + \varepsilon_t \quad (7)$$

Unlike synchronicity, increases in the values of *PrInfo* signify more informed trading. We therefore expect the signs on the coefficients of the explanatory variables to be the opposite of those reported in Table 3. The results reported in Panel B of Table 4 indeed appear strikingly similar, but with opposite signs on the coefficients. The coefficients on our key variable, *SentResPos*, are negative and statistically significant. This result indicates that informed trading becomes less prevalent during periods of high sentiment. The coefficients on *PeriodCount* and *PeriodCountSqr* are positive and negative, respectively, suggesting similar time trends in *PrInfo* as in stock return synchronicity. Overall, these results suggest that *PrInfo*, as an alternative measure of informed trading, also indicates that informed trading diminishes in importance during periods of high investor sentiment.

3.3. Cross-sectional variations in the relation between sentiment and return synchronicity

If the sentiment-synchronicity relationship is driven by time variations in the extent to which information is reflected in stock prices, we expect to see cross-sectional variations in this time-series relationship. As stocks that have more subjective valuations and are more difficult to arbitrage are more susceptible to prevailing sentiment, their increase in return synchronicity during extreme sentiment periods will be more pronounced as their stock prices are driven more by investor sentiment and less by firm-specific information. As pointed out by Baker and Wurgler (2006, 2007), Lemmon and Portniaguina (2006) and Qiu and Welch (2006), such a cross-sectional pattern is key to demonstrating that it is investor sentiment and limits to arbitrage that are at work.

To show that this cross-sectional variation in the time-series relationship is indeed present, we use two alternative approaches. The first approach is non-parametric. We perform a double sort on investor sentiment and stock characteristics and examine how stock return synchronicity varies along these two dimensions (Table 5, Panel A). We control for the potential effect of stock liquidity in a subsequent exercise (Table 5, Panels B and C), which sorts stocks further by their illiquidity (as measured by *PctZero*). The second approach makes use of panel regressions to examine how time-series variations in investor sentiment interact with cross-sectional stock characteristics in determining stock return synchronicity.

3.3.1. Cross-sectional variations in the sentiment-synchronicity link: Double sorts

We begin our investigation of cross-sectional variation in the sentiment-synchronicity relationship with a simple double sort. First, we examine whether synchronicity varies across months sorted on sentiment. Next, we examine if this time-series relationship varies in the cross-section across stock subgroups sorted by their characteristics.

Along the time-series dimension, we sort the 532 calendar months of our sample (October 1966 to January 2011) into quintiles based on the value of investor sentiment at the end of the previous month. Months with the lowest lagged sentiment values are placed in quintile 1. These are the periods with the most bearish sentiment. At the other extreme, months with the most bullish sentiment are placed in quintile 5. In contrast to these extremes, the middle quintiles represent periods in which sentiment is neither particularly bullish nor bearish. The effect of sentiment on the stock market is thus expected to be less pronounced during these periods. If sentiment-driven demand during positive sentiment periods is met with more limited arbitrage (due to short-sale constraints), less firm-specific information gets reflected in stock prices during these times and stock return synchronicity is expected to increase. This hypothesis predicts that synchronicity would be at its lowest for the middle quintiles and highest for quintile 5.

Along the cross-sectional dimension, we sort stocks into three subgroups based on each stock's composite characteristic rank (*Rank6V*) at the beginning of each month. As discussed in detail in Section 2.3, the composite rank of a stock is the first principle component of the decile ranks of the stock's size, age, stock return volatility, dividend payout, MB ratio and stock price level and is intended to capture a stock's sensitivity to investor sentiment. We label the group of stocks in the bottom (top) quintile of the composite rank as 'High Sensitivity' ('Low Sensitivity'). The remaining 60% of stocks are labelled 'Medium'.

For each firm in each month, we use the four-factor model given by Equation (1) to estimate return synchronicity. We still follow Equations (2) to (4) to aggregate the firm-level return synchronicity measures, only now the aggregation is performed at the level of each characteristic-sorted subgroup. We first examine how synchronicity varies over sentiment periods *within* each subgroup. We then calculate the difference in difference – the extent to which this sentiment-related difference in synchronicity is different *across* characteristic-sorted subgroups – to examine whether stocks that are being viewed as more sentiment-prone indeed experience larger increases in synchronicity when investor sentiment becomes more extreme. The results are reported in Table 5. The time-series variation in synchronicity over different sentiment quintiles within each subgroup is reported in Columns (2) to (4) of each panel. The difference in this time-series variation across subgroups is reported in Column (5).

We first examine how this sentiment-synchronicity relationship varies across different sensitivity rankings for all stocks. Table 5, Panel A reports these results. In Column (2), we see that for stocks that are more exposed to investor sentiment ('High Sensitivity' stocks), their average return synchronicity significantly increases from -0.349 to -0.228 as investor sentiment goes from quintile 3 (moderate) to quintile 5 (high). The difference of 0.121, reported in the last row, is statistically significant with a *t*-statistic of 3.33 (based on Newey-West standard errors). The difference is also economically significant, as the time series of monthly aggregate return synchronicity for this subgroup has a standard deviation of 0.201. In comparison, when sentiment moves from quintile 3 (moderate) to quintile 1 (low), average synchronicity increases but the difference of 0.029 is statistically insignificant. Thus, for stocks that are expected to be more exposed to sentiment a priori (i.e. 'High Sensitivity' stocks), we observe the same asymmetric effect of sentiment on stock return synchronicity that we report in Table 3.

In contrast, for stocks that are least exposed to investor sentiment ('Low Sensitivity' stocks), changes in their return synchronicity over time (as reported in Column (4) of Table 5, Panel A) do not exhibit such a relationship. This group of stocks represents companies that are large, mature, stable and high-dividend paying, with moderate valuations and high stock price levels. Baker and Wurgler (2007) argue that investor sentiment could affect such stocks in a fashion that is opposite to its effect on other stocks. In Column (5), we confirm that the sentiment-synchronicity relation is statistically different across stocks with different exposures to investor sentiment.

Next, we investigate if the pattern that we report in Table 5, Panel A is driven by stock liquidity. Every month, we sort stocks into two groups based on the percentage of their zero-return days (*PctZero*) that month. Stocks that are below the median *PctZero* for a given month are considered 'Liquid', and stocks that are above the median *PctZero* are considered 'Illiquid'.[15] We then repeat the analysis reported in Table 5, Panel A separately on the 'Liquid' and 'Illiquid' stocks.

Table 5, Panels B and C report these results. Our main findings in Table 5, Panel A – the sentiment-synchronicity relationship being asymmetric and cross-sectionally more pronounced for stocks that are more sensitive to sentiment – remain robust to variations in stock liquidity. The primary difference between Panels B and C is that the values of stock return synchronicity reported in Panel C are all lower than their counterparts in Panel B. This result is consistent with Chan, Hameed, and Kang (2013), who show that more illiquid stocks tend to experience lower stock return synchronicity.

3.3.2. Cross-sectional variations in the sentiment-synchronicity link: Multivariate regression analysis

Several other factors, besides investor sentiment, could influence the synchronicity of individual stock returns with the aggregate market. First, to ensure that our results are not driven by movements in interest rate and stock market liquidity, we use *SentRes* to measure variations in

[15] We obtain similar results when we assign the top and bottom 33% of the stocks by *PctZero* as 'Illiquid' and 'Liquid', respectively.

sentiment. To examine whether the cross-sectional variation in the relation between sentiment and return synchronicity continues to hold after controlling for such factors, we re-examine the above results in a panel regression framework. Our regression takes the following form:

$$Synchronicity_{j,t} = c_0 + c_1 SentResPos_{t-1} + c_2 AbsSentResNeg_{t-1} + c_3 SentResPos_{t-1} \times Rank6V_{j,t-1} + c_4 GDP_{t-1} + c_5 GDP_{t-1} \times Rank6V_{j,t-1} + c_6 Illiq_{t-1} + c_7 Illiq_{t-1} \times Rank6V_{j,t-1} + c_8 PeriodCount_t + c_9 PeriodCountSq_t + c_{10} ROESynchronicity_{j,t-1} + \varepsilon_{j,t} \quad (8)$$

As before, the subscript *j* indexes different stocks, and the subscript *t* indexes different months. The variable *Rank6V* is the decile rank of the stock based on the composite measure defined in Section 2.3. Recall that this measure is based on six individual characteristics, namely size, age, volatility, dividend-to-book ratio, MB ratio and stock price level, which are related to how sensitive a stock is to investor sentiment. We define this variable such that low ranks signify greater vulnerability to sentiment. The cross-sectional variation in the effect of high sentiment on stock return synchronicity is therefore captured by the coefficient of the interaction variable *SentResPos* x *Rank6V*. We expect the coefficient of the interaction variable, c_3, to be negative; the higher the rank of a stock, the weaker the effect of investor sentiment on its return synchronicity. Control variables are GDP growth rate (as motivated by the findings of Brockman, Liebenberg and Schutte 2010) and stock-level illiquidity (as motivated by Chan, Hameed and Kang 2013). We also allow the effect of these variables to vary in the cross-section, through their interaction terms with *Rank6V*. We compute two stock-level measures of illiquidity and include them alternately as controls in all of our cross-sectional analyses. The first is the percentage of zero-return days

(*PctZero*) in a given month, as introduced by Lesmond, Ogden and Trzcinka (1999). The second is the average daily Amihud (2002) illiquidity over a month, in which daily illiquidity of a stock is computed as the absolute daily return on the stock divided by its daily dollar volume. We compute these measures of illiquidity for each stock in each month. Higher values of these measures indicate greater illiquidity.

We estimate Equation (7) using pooled, stock-month data. The results are reported in Table 6. We first estimate the model without the illiquidity control variables (reported in Column (1)), and then add the alternate measures of illiquidity back in Columns (2) and (3). The reported *t*-statistics are based on standard errors clustered by both firm and month. Coefficients on *SentResPos* are positive and statistically significant in all specifications, but the coefficients on *AbsSentResNeg* are not. This result indicates that the asymmetric effect of sentiment on stock return synchronicity holds even after controlling for a host of other time-series and cross-sectional variables. In addition, the interaction term *SentResPos* x *Rank6V* is negative, as hypothesised, and is highly statistically significant. This finding confirms that the asymmetric effect of sentiment on stock return synchronicity exhibits a cross-sectional pattern – stronger for stocks with a lower *Rank6V* and hence a higher sensitivity to sentiment. This result also shows that the findings we document in a simpler, bivariate framework in Table 5 carry over to a multiple regression setting.

Although the coefficients on standalone *GDP* are positive and only marginally significant (with point estimates around 0.02), the interaction term, *GDP* x *Rank6V*, is negative and highly significant (with point estimates around -0.01). As *Rank6V* ranges from 1 to 10 with both mean and median around 5, our panel regression results suggest that the relationship between *GDP* and synchronicity for the average firm is negative, consistent with the results obtained by Brockman, Liebenberg and Schutte (2010). More notable, however, is our finding that the negative relation between real GDP growth and return synchronicity varies in the cross-section, as reflected by the negative and significant coefficient on the interaction term. This finding is noteworthy for two reasons. First, it offers novel evidence that the rational channel of time-varying return synchronicity, proposed by Veldkamp (2005, 2006) and empirically supported by Brockman et al. (2010), exhibits a cross-sectional pattern that is stronger for easier-to-value companies (i.e. firms that are large, mature, pay high dividends, and have moderate valuations and high stock prices). Second, because we find earlier that the time-series relation between sentiment and return synchronicity also exhibits a cross-sectional pattern – but one that *weakens* (rather than strengthens) for easier-to-value stocks – our results here further demonstrate that aggregate investor sentiment affects return synchronicity through a behavioural channel that is quite distinct from the rational one examined by Veldkamp (2005, 2006) and Brockman et al. (2010).

Although the coefficients on the standalone illiquidity measures (*PctZero* or *Illiq*) are positive (with point estimates around 0.001), the interaction terms, *PctZero* x *Rank6V* or *Illiq* x *Rank6V*, are negative and highly significant (with point estimates around -0.001). As *Rank6V* ranges from 1 to 10 with both mean and median around 5, our panel regression results suggest that the relationship between the illiquidity measures and synchronicity for the average firm is negative, consistent with the findings of Chan, Hameed and Kang (2013). The negative sign on the interaction terms (*PctZero* x *Rank6V* or *Illiq* x *Rank6V*) in Columns (2) and (3) indicates that the negative relationship between illiquidity and stock return synchronicity varies in the cross-section, becoming more pronounced for stocks that are easier to value.

We also find that the coefficients for standalone *Rank6V* are positive and significant. This finding is consistent with the results reported by Cao, Simin and Zhao (2008) and Brandt, Brav, Graham and Kumar (2010), who show that small and low-priced stocks tend to have lower return synchronicity than large and high-priced stocks. Finally, we note that the coefficient on *ROE Synchronicity* is positive and highly statistically significant. This is consistent with the evidence presented by Wei and Zhang (2006), Irvine and Pontiff (2009) and Bekaert, Hodrick and Zhang (2012) and indicates that the synchronicity of fundamentals is an important driver of the synchronicity of stock returns.

Overall, our panel regression results are consistent with the non-parametric, double-sort results reported in Table 5. The relation between sentiment and return synchronicity is asymmetric and cross-sectionally more pronounced for those stocks that previous studies suggest should have a higher exposure to aggregate investor sentiment.

4. Robustness checks

We carry out two sets of robustness checks on our main findings. Our first set relates to the relation between sentiment and aggregate market-wide synchronicity reported in Table 3. We add additional market-wide controls to see if the results still hold and also examine the effect of an alternative weighting scheme in computing aggregate market-wide synchronicity. Our second set of robustness checks relate to the panel regressions in Table 6. These include examining if our findings are robust to the use of an alternative sentiment index that is constructed recursively, a PLS-based investor sentiment index constructed by Huang, Jiang, Tu and Zhou (2015), a sentiment measure that is contemporaneous to return synchronicity and that excludes the Internet bubble period of the late 1990s.

4.1. DOX and market volatility as additional controls in aggregate regressions

Unlike the Baker-Wurgler sentiment index, which is constructed using a top-down approach, a recent paper by Cassella and Gulen (2018) proposes a market-wide measure of investors' behavioural biases that is constructed using a bottom-up approach. Named 'the degree of extrapolation bias' or DOX, it uses survey expectations data to capture the tendency of investors to extrapolate recent returns when forming beliefs about future returns. To examine how synchronicity is related to this alternative measure of investors' psychological biases, we use the extended version of *DOX* computed by Cassella and Gulen (2018), which begins in December 1967.[16] Higher values of the index signify periods when investors rely more on recent returns in forming their expectations. From Table 2, we see that *DOX* is positively correlated with the two sentiment indices (*Sent* and *Sent_PLS*) but only moderately so, with correlation coefficients of less than 0.5. As Cassella and Gulen (2018) show that *DOX* is a positive and significant predictor of aggregate stock return volatility, we examine whether *DOX* also drives the relationship between *SentResPos* and return synchronicity. Cassella and Gulen (2018, Table 8) show that *DOX* is a significant standalone predictor of aggregate volatility. Consistent with this result, we also find in Table 7, Column (1) that standalone *DOX* – but not its interaction with *D/P* – has significant predictive power for return synchronicity. In Table 7, Column (2), we see that even after *DOX* has been controlled for, the coefficient on *SentResPos* remains positive and significant.

Bekaert, Hodrick and Zhang (2009) and others show that stock return comovement strengthens mechanically in times of high market volatility. If sentiment is correlated with market

[16] We extract the approximate values of *DOX* from Figure A4 of Cassella and Gulen (2018) using a digitizing software, Getdata Graph Digitizer, and linear interpolation.

volatility, our results may simply reflect the positive association between market volatility and synchronicity. To examine this possibility, we include market volatility as an additional control in Table 7, Column 3. Consistent with prior literature, the coefficient on market volatility is positive and highly significant, indicating that stock return synchronicity indeed increases with market volatility. However, the coefficients on the two sentiment variables indicate that the inclusion of market volatility as an additional control has little effect on the relationship between sentiment and stock return synchronicity. From Table 2, we see that the Baker-Wurgler sentiment index and market volatility are only weakly correlated, with a correlation coefficient of 0.08.

4.2. Market-value weighting

As displayed in Equation (3) above, the way we construct aggregate R^2 follows MYY, by weighting firm-level R^2 by each firm's sum of squared total variations (*SST*). In Table 8, Column (4), we use each firm's market value of equity as alternative weights. The coefficient estimates become insignificant, confirming that the effect of sentiment on stock return synchronicity is indeed more pronounced for smaller firms.

4.3. Recursive sentiment index

The Baker-Wurgler sentiment index is estimated using the full sample and, hence, suffers from look-ahead biases. In this section, we follow the recursive procedure outlined in Huang, Jiang, Tu and Zhou (2015, pp. 807-809) to construct an alternative index that is free from look-ahead biases. We first set the July 1965 to December 1984 period as our initial estimation sample.

Applying the Baker and Wurgler (2006, 2007) method to the initial estimation sample allows us to obtain a December 1984 value for investor sentiment that does not rely on any future observations. For all subsequent months, we recursively obtain estimates of investor sentiment in this way, using only data already known up to that point in time. This procedure yields monthly values of investor sentiment from December 1984 through December 2010. This time series of recursively estimated sentiment, denoted *RecSent*, is free of look-ahead biases. Next, we orthogonalise *RecSent* with respect to the short-term interest rate *Tbill* and stock market illiquidity *PctZero* to obtain *RecSentRes* and repeat our panel regression analysis. Table 8, Column (1) reports these results, which show that our main results are robust: the sentiment-synchronicity relationship remains asymmetric and cross-sectionally more pronounced for those stocks that are more sensitive to investor sentiment (i.e. with low *Rank6V*).

4.4. Huang et al. (2015) investor sentiment index

The Baker-Wurgler index is based on the first principal component of six individual sentiment proxies. Huang, Jiang, Tu and Zhou (2015) use PLS to extract an alternative sentiment index from the same underlying individual proxies and show that this index has strong predictive power for aggregate stock market returns. We report results in Table 8, Column (2) that use this PLS-based index as an alternative investor sentiment measure. We show that, even when the PLS-based index is used, the sentiment-synchronicity relationship remains asymmetric and cross-sectionally more pronounced for those stocks that are more sensitive to investor sentiment.[17]

4.5. Removing the Internet bubble period

The Internet bubble period of the late 1990s was a period of extreme exuberance that lasted from 1997 to early 2000. It is possible that our main finding of a link between synchronicity and high sentiment is primarily driven by this period. In Table 8, Column (3), we replicate the results in Table 6 after excluding the January 1997 to March 2000 period. We find that our conclusion is not sensitive to the exclusion of this period.

4.6. Contemporaneous sentiment

Our analysis so far examines the relationship between investor sentiment in month t-1 and return synchronicity in month t. In this section, we explore the robustness of our findings to the use of contemporaneous sentiment and synchronicity, matching investor sentiment with stock

---

[17] Jiang, Lee, Martin, and Zhou (2018) develop a manager sentiment index based on the aggregated textual tone of corporate financial disclosures. Unlike the Baker-Wurgler and Huang, Jiang, Tu, and Zhou (2015) PLS indices, Jiang, Lee, Martin, and Zhou's (2018) manager sentiment index is available for a much shorter period, beginning only in 2003. In untabulated results, we find that the manager sentiment index is highly correlated with *Sent*, with a correlation coefficient of 0.793. For our aggregate market-wide analysis in Table 3, we obtain qualitatively similar results using Jiang, Lee, Martin, and Zhou's (2018) manager sentiment index. However, for the panel regressions in Table 6, the results become weak using this alternative index.

return synchronicity in the same month. These results are reported in Table 7, Column (4). We find that our conclusion is not sensitive to this alternative timing convention.

5. Conclusion

While previous studies focus on the effect of aggregate investor sentiment on the levels of asset prices and returns, we extend the literature and examine the relationship between aggregate sentiment and stock return synchronicity. From a second-moment perspective that is distinct from the focus of prior works, our evidence provides independent support that aggregate investor sentiment affects asset prices. This is consistent with the previous finding that mispricing is more serious when sentiment is strongest and for stocks most affected by sentiment. We show that it is precisely during those times and for those stocks that return synchronicity is highest, consistent with sentiment-driven trading being more prevalent and arbitrage constraints more binding in such cases.

The study of stock return synchronicity is of fundamental concern to investors and portfolio managers. The extent to which different stocks move together affects the benefit of portfolio diversification and the potential payoff to stock selection. Our contribution is twofold. First, we show that stock return synchronicity varies over time as a function of aggregate investor sentiment. Second, we show that there are cross-sectional variations in this time-series relationship. Our results show that the variations along both of these dimensions are significant.

Understanding the circumstances under which stock return synchronicity moves has additional implications. Prior work shows that higher stock return synchronicity is associated with corporate managers being less reliant on stock prices in their investment decisions (Chen, Goldstein, and Jiang 2007). Our evidence that synchronicity increases during periods of bullish sentiment is therefore relevant for corporate managers who intend to obtain feedback on the quality of their decisions from stock prices. Prior studies also show that more synchronous stock prices are associated with a less efficient allocation of resources in the economy (Wurgler 2000 and Durnev, Morck, and Yeung 2004). Our findings are thus of interest to policy makers as it underscores a distinct channel through which investor sentiment can impede the capital allocation role of financial markets. We believe future research would continue to deepen our understanding of the effects of investor sentiment on financial markets.

Acknowledgements

We thank Geert Bekaert (the editor), Kalok Chan, Richard Chung, Allaudeen Hameed, Chuan Yang Hwang, Marcin Kacperczyk, Alok Kumar, Srinivasan Sangakaraguruswamy, and three anonymous reviewers for helpful comments, Manhong Chan and Roy Chu for excellent research assistance, and the General Research Fund of the Research Grants Council of the Hong Kong Special Administrative Region, China, (Project No. 590313) and the Hong Kong Polytechnic University for research support. We remain responsible for any errors.

References

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

Ang, A., R.J. Hodrick, Y. Xing, and X. Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259-299.

Antoniou, C., J.A. Doukas, and A. Subrahmanyam, 2013, Cognitive dissonance, sentiment, and momentum, *Journal of Financial and Quantitative Analysis* 48, 245-275.

Antoniou, C., J.A. Doukas, and A. Subrahmanyam, 2015, Investor sentiment, beta, and the cost of equity capital, *Management Science* 62, 347-367.

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

Baker, M. and J. Wurgler, 2007, Investor sentiment in the stock market, *Journal of Economic Perspectives* 21, 129-151.

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

Ball, R., G. Sadka, and R. Sadka, 2009, Aggregate earnings and asset prices, *Journal of Accounting Research* 47, 1097-1133.

Barberis, N., 2018, Psychology-based models of asset prices and trading volume, working paper, Yale University.

Barberis, N. and A. Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 161-199.

Barberis, N., A. Shleifer, and J. Wurgler, 2005, Comovement, *Journal of Financial Economics* 75, 283-317.

Bartram, S., G. Brown, and R. Stulz, 2012, Why are U.S. stocks more volatile?, *Journal of Finance* 67, 1329-1370.

Bekaert, G. and R. Hodrick, and X. Zhang, 2009, International stock return comovements, *Journal of Finance* 64, 2591-2626.

Bekaert, G. and R. Hodrick, and X. Zhang, 2012, Aggregate idiosyncratic volatility, *Journal of Financial and Quantitative Analysis* 47, 1155-1185.

Bekaert, G. and M. Hoerova, 2016, What do asset prices have to say about risk appetite and uncertainty? *Journal of Banking and Finance* 67, 103-118.

Boubaker, S., H. Mansali, and H. Rjiba, 2014, Large controlling shareholders and stock price synchronicity, *Journal of Banking and Finance* 40, 80-96.

Brandt, M., A. Brav, J.R. Graham, and A. Kumar, 2010, The idiosyncratic volatility puzzle: Time trend or speculative episodes, *Review of Financial Studies* 23, 863-899.

Bris, A., W. Goetzmann, and N. Zhu, 2007, Efficiency and the bear: Short-sales and markets around the world, *Journal of Finance* 62, 1029-1079.

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

Brockman, P. and X. Yan, 2009, Block ownership and firm-specific information, *Journal of Banking and Finance* 33, 308-316.

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

Cao, C., T. Simin, and J. Zhao, 2008, Can growth options explain the trend in idiosyncratic risk? *Review of Financial Studies* 21, 2599-2633.

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

Cassella S. and H. Gulen, 2018, Extrapolation bias and the predictability of stock returns by price-scaled variables, *Review of Financial Studies* 31, 4345-4397.

Chan, K. and A. Hameed, 2006, Stock price synchronicity and analyst coverage in emerging markets, *Journal of Financial Economics* 80, 115-147.

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

Chen, Q., I. Goldstein, and W. Jiang, 2007, Price informativeness and investment sensitivity to stock prices, *Review of Financial Studies* 20,

Cooper. M.J., O. Dimitrov, and P.R. Rau, 2001, A Rose.com by any other name, *Journal of Finance* 56, 2371-2388.

Cooper. M.J., A. Khorana, I. Osobov, A. Patel, and P.R. Rau, 2005, Managerial actions in response to a market downturn: Valuation effects of name changes in the dot.com decline, *Journal of Corporate Finance* 11, 319-335.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam, 1998, Investor Psychology and Security Market Under- and Overreactions, *Journal of Finance* 53, 1839-1885.

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

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

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

Engle, R., 2002, Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroscedasticity models, *Journal of Business & Economic Statistics* 20, 339-350.

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

Fama, E.F. and K.R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1-22.

Fernandes, N. and M. Ferreira, 2008, Does international cross-listing improve the information environment?, *Journal of Financial Economics* 88, 216-244.

Fernandes, N. and M. Ferreira, 2009, Insider trading laws and stock price informativeness, *Review of Financial Studies* 22, 1845-1887.

Ferreira, D., M. Ferreira, and C. Raposo, 2011, Board structure and price informativeness, *Journal of Financial Economics* 99, 523-545.

Ferreira, M. and P. Laux, 2007, Corporate governance, idiosyncratic risk, and information flow, *Journal of Finance* 62, 951-989.

Francis, J., S. Huang, I. Khurana, and R. Pereira, 2009, Does corporate transparency contribute to efficient resource allocation?, *Journal of Accounting Research* 47, 943-989.

Greenwood, R. and S. Hanson, 2013, Issue quality and corporate bond returns, *Review of Financial Studies* 26, 1483-1525.

Gul, F., J. Kim, and A. Qiu, 2010, Ownership concentration, foreign shareholding, audit quality, and stock price synchronicity: Evidence from China, *Journal of Financial Economics* 95, 425-442.

Gul, F., B. Srinidhi, and A. Ng, 2011, Does board gender diversity improve the informativeness of stock prices?, *Journal of Accounting and Economics* 51, 314-338.

Huang, D., F. Jiang, J. Tu, and G. Zhou, 2015, Investor sentiment aligned: A powerful predictor of stock returns, *Review of Financial Studies* 28, 791-837.

Irvine, P. and J. Pontiff, 2009, Idiosyncratic return volatility, cash flows, and product market competition, *Review of Financial Studies* 22, 1149-1177.

Jiang, F., J. Lee, X. Martin, and G. Zhou, 2018, Manager sentiment and stock returns, *Journal of Financial Economics*, forthcoming.

Jin, L. and S. Myers, 2006, R^2 around the world: New theory and new tests, *Journal of Financial Economics* 79, 257-292.

Kim, Y., H. Li, and S. Li, 2012, Does eliminating the Form 20-F reconciliation from IFRS to U.S. GAAP have capital market consequences?, *Journal of Accounting and Economics* 53,249-270.

Kurov, A., 2010, Investor sentiment and the stock market's reaction to monetary policy, *Journal of Banking and Finance* 34, 139-149

Lee, K., 2011, The world price of liquidity risk, *Journal of Financial Economics* 99, 136-161.

Lemmon, M. and E. Portniaguina, 2006, Consumer confidence and asset prices: Some empirical evidence, *Review of Financial Studies* 19, 1499-1529.

Lesmond, D., J. Ogden, and C. Trzcinka, 1999, A new estimate of transaction costs, *Review of Financial Studies* 12, 1113-1141.

Llorente, G., R. Michaely, G. Saar, and J. Wang, 2002, Dynamic volume-return relation of individual stocks, *Review of Financial Studies* 15, 1005-1047.

Mian, G.M. and S. Sankaraguruswamy, 2012, Investor sentiment and stock market response to earnings news, *The Accounting Review* 87, 1357-1384.

Morck, R., B. Yeung, and W. Yu, 2000, The information content of stock markets: Why do emerging markets have synchronous price movements? *Journal of Financial Economics* 58, 215-260.

Newey, W. and K. West, 1987, A simple, positive semi-definite, heteroscedastic and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.

Piotroski, J. and D. Roulstone, 2004, The influence of analysts, institutional investors, and insiders on the incorporation of market, industry, and firm-specific information into stock prices, *The Accounting Review* 79, 1119-1151.

Qiu L. and I. Welch, 2006, Investor sentiment measures, SSRN working paper. Available at: https://ssrn.com/abstract=595193.

Rajgopal, S. and M. Venkatachalam, 2011, Financial reporting quality and idiosyncratic return volatility, *Journal of Accounting and Economics* 51, 1-20.

Riedl, E. and G. Serafeim, 2011, Information risk and fair values: An examination of equity betas, *Journal of Accounting Research* 49, 1083-1122.

Shleifer, A. and R.W. Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52, 35-55.

Sibley, S.E., Y. Wang, Y. Xing, X. Zhang, 2016, The information content of the sentiment index, *Journal of Banking & Finance* 62, 164-179.

Stambaugh, R.F., J. Yu, and Y. Yuan, 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104, 288-302.

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

Veldkamp, L., 2005, Slow boom, sudden crash, *Journal of Economic Theory* 124, 230-257.

Veldkamp, L., 2006, Information markets and the comovement of asset prices, *Review of Economic Studies* 73, 823-845.

Wei, S. and C. Zhang, 2006, Why did individual stocks become more volatile? *Journal of Business* 79, 259-292.

Wurgler, J., 2000, Financial markets and the allocation of capital, *Journal of Financial Economics* 58, 187-214.

Xu, N., K.C. Chan, X. Jiang, and Z. Yi, 2013, Do star analysts know more firm-specific information? Evidence from China, *Journal of Banking and Finance* 37, 89-102.

Yu, J. and Y. Yuan, 2011, Investor sentiment and the mean-variance relation, *Journal of Financial Economics* 100, 367-381.

Table 1
Descriptive statistics

This table reports summary statistics for the key variables in our analyses. All variables are defined in Appendix A.

	Variable	No. of Observations	Mean	Median	Standard Deviation	Minimum	Maximum
Panel A: Measure of synchronicity (stock-month and monthly observations from February 1966 to January 2011)							
MYY Synchronicity (stock-month observations)	$MYY\_Synchronicity_{j,t}$	1,902,193	-1.097	-1.089	1.054	-3.825	1.376
MYY Synchronicity (monthly time series)	$MYY\_Synchronicity_{t}$	540	-1.120	-1.162	0.269	-1.660	0.172
Fama-French-Carhart 4-factor Synchronicity (stock-month observations)	$FF4F\_Synchronicity_{j,t}$	1,902,193	-0.135	-0.153	0.803	-2.044	1.856
Fama-French-Carhart 4-factor Synchronicity (monthly time series)	$FF4F\_Synchronicity_{t}$	540	-0.242	-0.273	0.212	-0.745	0.889
Panel B: Measure of investor sentiment (monthly observations from January 1966 to December 2010)							
Baker-Wurgler Investor Sentiment Index	*Sent*	540	0.019	-0.011	0.988	-2.578	2.691
SentRes (residuals, based on the 2-variable model)	*SentRes*	540	0.017	-0.056	0.792	-1.835	1.970
PLS Sentiment Index	*Sent_PLS*	540	0.018	-0.258	0.990	-1.724	3.196
Degree of Extrapolative Weighting (from December 1967 to December 2010)	*DOX*	517	0.424	0.384	0.194	0.122	0.902
Panel C: Control variables (monthly observations from January 1966 to December 2010)							
3-month T-bill Rate (%)	*Tbill*	540	5.509	5.185	2.967	0.030	16.300
Proportion of Zero Return (%)	*PctZero*	540	24.983	24.937	14.395	2.312	47.634
Dividend-price Ratio (%)	*D/P*	540	0.252	0.197	0.160	0.067	0.870
Market Volatility (%)	*MktVol*	540	13.572	11.147	8.197	3.954	78.465
Panel D: Control variables (quarterly observations from 1966 Q1 to 2010 Q4)							
Real GDP Growth (%)	*GDP*	180	0.709	0.747	0.861	-2.301	3.934
Synchronicity of ROE (from 1966 Q3 to 2010 Q4)	*ROE Synchronicity*	178	-0.857	-1.068	0.727	-1.758	3.540
Panel E: Control variables (stock-month observations from February 1966 to January 2011)							
Synchronicity of ROE (from 1966 Q3 to 2010 Q4)	*ROE Synchronicity*	1,704,742	-1.047	-0.996	1.696	-13.445	23.227
Illiquidity (x 1,000,000)	*Illiq*	1,902,193	4.859	0.153	17.910	0.000	143.664
Proportion of Zero Return (%)	*PctZero*	1,902,193	19.519	14.286	19.330	0.000	100.000
Rank of Firm Size (1 - smallest; 10 - largest)	*RankSize*	1,902,193	3.215	2.000	2.789	1.000	10.000
Rank of Firm Age (1 - youngest; 10 - oldest)	*RankAge*	1,902,193	3.728	3.000	2.556	1.000	10.000
Rank of Stock Return Volatility (1 - most volatile; 10 - least volatile)	*RankStd*	1,902,193	4.205	3.000	3.018	1.000	10.000

Rank of Dividend-to-Book Ratio (0 - non-payers; 1 - lowest non-zero ratio; 10 - highest ratio)	*RankDiv*	1,902,193	2.353	0.000	3.210	0.000	10.000
Rank of MB Ratio (1.5 - extreme ratio; 9.5 - medium ratio)	*RankMB*	1,902,193	5.000	5.500	2.884	1.500	9.500
Rank of Stock Price (1 - lowest; 10 - highest)	*RankPrc*	1,902,193	3.709	3.000	2.859	1.000	10.000
Combined Rank of the 6 Rank Variables (1 - lowest; 10 - highest)	*Rank6V*	1,902,193	5.493	5.000	2.880	1.000	10.000

Table 2
Correlation matrix

The Pearson correlation coefficients reported in this table are based on time-series data over the February 1966 to January 2011 period (540 months), except for *Sent_JLMZ* and *DOX* that are available for shorter periods. *MYY_Synchronicity* is Morck, Yeung and Yu's (2000) measure of stock return synchronicity. *FF4F_Synchronicity* and *FF5F_Synchronicity* use, respectively, the Fama-French-Carhart four-factor model and the Fama-French (2015) five-factor model to compute stock return synchronicity. *Sent* is the Baker and Wurgler (2006, 2007) monthly sentiment index. *SentRes* is obtained by orthogonalising the Baker-Wurgler index with respect to the T-bill rate and Lee's (2011) liquidity factor. *Sent_PLS* is the PLS-based sentiment index constructed by Huang, Jiang, Tu and Zhou (2015). *MktVol* is market volatility computed as the annualised standard deviation of market daily return within each month. *DOX* is the Cassella and Gulen (2018) index that captures the degree of investors' extrapolative beliefs. *Tbill* is the three-month T-bill rate. *PctZero* is the aggregated market-level measure (by taking a simple average over individual stocks) of the per cent of zero-return days in a month. *GDP* is the growth rate in real gross domestic product. *ROE Synchronicity* is the synchronicity of quarterly ROE, estimated from the most recent 12 quarters. The *t*-statistics are reported in parentheses and are computed based on Newey-West (1987) standard errors.

	MYY_ Synchronicity	*FF4F_ Synchronicity*	*FF5F_ Synchronicity*	*Sent*	*SentRes*	*Sent_PLS*	*Sent_JLMZ*	*DOX*	*Tbill*	*PctZero*	*GDP*	*MktVol (%)*	*ROE Synchronicity*
MYY_Synchronicity	1.000	0.940 (< 0.001)	0.884 (< 0.001)	0.149 (0.001)	-0.083 (0.054)	0.289 (< 0.001)	0.325 (0.001)	0.379 (< 0.001)	-0.049 (0.254)	-0.341 (< 0.001)	-0.164 (< 0.001)	0.286 (< 0.001)	0.248 (< 0.001)
FF4F_Synchronicity		1.000	0.958 (< 0.001)	0.142 (0.001)	-0.061 (0.154)	0.256 (< 0.001)	0.266 (0.009)	0.309 (< 0.001)	-0.037 (0.391)	-0.295 (< 0.001)	-0.135 (0.002)	0.268 (< 0.001)	0.241 (< 0.001)
FF5F_Synchronicity			1.000	0.133 (0.002)	-0.050 (0.250)	0.227 (< 0.001)	0.270 (0.008)	0.256 (< 0.001)	-0.046 (0.290)	-0.281 (< 0.001)	-0.103 (0.017)	0.220 (< 0.001)	0.211 (< 0.001)
Sent				1.000	0.796 (< 0.001)	0.726 (< 0.001)	0.793 (<0.0001)	0.338 (< 0.001)	0.272 (< 0.001)	-0.193 (< 0.001)	-0.089 (0.039)	0.076 (0.079)	0.039 (0.370)
SentRes					1.000	0.533 (< 0.001)	0.023 (0.825)	0.012 (0.780)	-0.011 (0.806)	-0.004 (0.923)	-0.030 (0.487)	0.029 (0.508)	-0.058 (0.179)
Sent_PLS						1.000	0.019 (0.856)	0.312 (< 0.001)	0.274 (< 0.001)	-0.107 (0.013)	-0.027 (0.535)	0.150 (0.001)	0.149 (0.001)
Sent_JLMZ							1.000	0.418 (< 0.001)	0.405 (< 0.001)	-0.629 (< 0.001)	-0.064 (0.534)	0.081 (0.432)	-0.161 (0.118)
DOX								1.000	-0.010 (0.814)	-0.440 (< 0.001)	-0.325 (< 0.001)	0.341 (< 0.001)	-0.058 (0.189)
Tbill									1.000	0.696 (< 0.001)	-0.013 (0.766)	-0.179 (< 0.001)	0.074 (0.088)
PctZero										1.000	0.073 (0.091)	-0.263 (< 0.001)	-0.033 (0.450)
GDP											1.000	-0.307 (< 0.001)	0.013 (0.771)
MktVol (%)												1.000	-0.093 (0.032)
ROE Synchronicity													1.000

Table 3
Investor sentiment and aggregate return synchronicity

This table reports multiple regressions of aggregate stock return synchronicity on investor sentiment and other control variables. *SentPos* takes the value of the Baker-Wurgler sentiment index (*Sent*) when it is positive and zero otherwise. *AbsSentNeg* takes the absolute value of *Sent* when it is negative and zero otherwise. *SentResPos* takes the value of the orthogonalised Baker-Wurgler sentiment index (*SentRes*) when it is positive and zero otherwise. *AbsSentResPos* takes the absolute value of *SentRes* when it is negative and zero otherwise. *SentRes* represents the Baker-Wurgler sentiment index that has been orthogonalised with respect to the T-bill rate and Lee's (2011) liquidity factor. *GDP* is the growth rate in real gross domestic product. *PeriodCount* and *PeriodCountSq* represent linear and quadratic time effects, respectively. *ROE Synchronicity* is the synchronicity of quarterly ROE, estimated from the most recent 12 quarters. In specifications (1), Morck, Yeung and Yu's (2000) measure of stock return synchronicity is used. Specification (2) and (3) use *FF4F_Synchronicity*, which relies on the Fama-French-Carhart four-factor model. Specification (4) is based on *FF5F-Synchronicty*, which uses the Fama-French (2015) five-factor model. Specification (5) uses *FF4GF_Synchronicity*, which is analogous to *FF4F_Synchronicity* except that it uses global versions of the market, size, value and momentum factors from 1991 onwards. *D/P* is the aggregate dividend-price ratio. All results are obtained using monthly observations. The sample period is from October 1966 to January 2011. The *t*-statistics are reported in parentheses and are based on Newey-West (1987) standard errors.

	Dependent variable				
	MYY_ Synchronicity	*FF4F_Synchronicity*		*FF5F_ Synchronicity*	*FF4GF_ Synchronicity*
Explanatory variables	(1)	(2)	(3)	(4)	(5)
Intercept	-0.755	0.124	0.088	0.382	0.073
	(-9.12)	(1.82)	(1.24)	(5.59)	(1.06)
SentPos	0.112	0.082			
	(4.26)	(3.93)			
AbsSentNeg	0.004	-0.025			
	(0.14)	(-0.92)			
SentResPos			0.089	0.065	0.090
			(3.48)	(2.97)	(3.55)
AbsSentResNeg			0.046	0.042	0.045
			(1.09)	(1.08)	(1.09)
GDP	-0.051	-0.022	-0.024	-0.016	-0.019
	(-3.63)	(-2.09)	(-2.11)	(-1.42)	(-1.87)
PeriodCount	-0.004	-0.003	-0.003	-0.002	-0.003
	(-5.84)	(-5.67)	(-5.32)	(-4.74)	(-5.04)
PeriodCountSqr	0.00001	0.00001	0.00001	0.000005	0.000005
	(7.02)	(6.82)	(6.18)	(5.62)	(5.64)
ROE Synchronicity	0.006	0.024	0.034	0.011	0.036
	(0.21)	(0.59)	(1.05)	(0.40)	(1.09)
Adj R^2	0.29	0.31	0.28	0.19	0.27
No. of months	532	532	532	532	532

Table 4
Private-information-based trading and investor sentiment

Panel A reports the summary statistics of the Llorente et al. (2002) measure of private-information-based-trading (*PrInfo*). We first compute *PrInfo* for each stock in each month. We then obtain an aggregate *PrInfo* measure for each month by taking an equal-weighted average of all *PrInfo* at the individual stock level. Panel B reports the results of the regression of aggregate *PrInfo* on investor sentiment and controls. The sentiment index used, *SentRes*, is the Baker-Wurgler index orthogonalised to the T-bill rate and Lee's (2011) liquidity factor. *SentResPos* takes the value of *SentRes* when it is positive and zero otherwise. *AbsSentResNeg* takes the absolute value of *SentRes* when it is negative and zero otherwise. *GDP* is the growth rate in real gross domestic product. *PeriodCount* and *PeriodCountSq* represent linear and quadratic time effects, respectively. *ROE Synchronicity* is the synchronicity of quarterly ROE, estimated from the most recent 12 quarters. The regressions are estimated using monthly observations from October 1966 to January 2011. The *t*-statistics are reported in parentheses and are based on Newey-West (1987) standard errors.

Panel A: Summary statistics of the private-information-based trading (*PrInfo*; x100)

	No. of Obs.	Mean	Median	Std. Dev.	Minimum	Maximum
Stock-month observations	2,234,569	-2.160	-0.503	37.290	-122.364	109.425
Monthly time series	532	-2.583	-2.323	3.193	-16.507	13.337

Panel B: Regression of *PrInfo* on sentiment and control variables

	Dep var = *PrInfo*	
Explanatory variables	(1)	(2)
Intercept	-7.144	-7.423
	(-3.31)	(-2.92)
SentResPos	-2.950	-2.953
	(-2.88)	(-2.90)
AbsSentResNeg	-1.285	-1.229
	(-1.44)	(-1.19)
GDP	0.058	0.060
	(0.22)	(0.23)
PeriodCount	0.048	0.052
	(3.15)	(2.26)
PeriodCountSqr	-0.0001	-0.0001
	(-3.14)	(-2.40)
ROE Synchronicity		0.302
		(0.23)
Adj R^2	0.09	0.09
No. of months	532	532

Table 5
Cross-sectional variation in the sentiment-synchronicity relation: Bivariate sorts

This table reports the cross-sectional variation in the relation between investor sentiment and stock return synchronicity. At the beginning of each calendar month, we sort all stocks based on a composite rank that captures the sensitivity of the stock to investor sentiment. The composite rank is based on the first principal component of the ranks of a stock on six stock characteristics, namely size, age, stock return volatility, dividend payout, the MB ratio and the stock price level. Stocks that fall in the two extreme quintiles are labelled as 'High Sensitivity' and 'Low Sensitivity', and the remaining 60% of stocks are classified as 'Medium'. Within each group, we then compute average synchronicity within a month. Months are then sorted into quintiles based on the value of the Baker-Wurgler investor sentiment index at the end of the previous month. Finally, we compute average synchronicity for each subgroup within a sentiment quintile by taking a simple average of all of the synchronicity measures for that subgroup that fall in that sentiment quintile. The row 'Difference (Low - Moderate)' reports the difference in stock return synchronicity between the 'Low' and 'Moderate' sentiment quintiles. The row 'Difference (High - Moderate)' reports the difference in stock return synchronicity between the 'High' and 'Moderate' sentiment quintiles. Panel A reports the analysis for the total sample. Panels B and C repeat the same analysis for subsamples of liquid and illiquid stocks, respectively. At the end of every month, all stocks are ranked based on their *PctZero* – the proportion of the stocks' zero-return days in a month. Stocks that are below the median *PctZero* for a given month are considered 'Liquid', and stocks that are above the median *PctZero* are considered 'Illiquid' The analysis is based on monthly data from October 1966 to January 2011. The *t*-statistics are reported in parentheses and are based on Newey-West (1987) standard errors.

Panel A: All stocks

	Average monthly synchronicity			Cross-sectional difference
Quintiles of sentiment	High sensitivity	Medium	Low sensitivity	
(1)	(2)	(3)	(4)	(5) = (2) - (4)
1 (Low)	-0.319	-0.169	0.070	
2	-0.351	-0.232	0.154	
3 (Moderate)	-0.349	-0.218	0.132	
4	-0.317	-0.173	0.210	
5 (High)	-0.228	-0.121	0.087	
Difference (Low - Moderate)	0.029	0.049	-0.062	0.092
t-statistics	(0.96)	(1.13)	(-1.05)	(1.85)
Difference (High - Moderate)	0.121	0.098	-0.045	0.166
t-statistics	(3.33)	(2.08)	(-0.78)	(3.20)

Table 5 (Continued)

Panel B: Liquid stocks

Quintiles of sentiment	Average monthly synchronicity High sensitivity	Medium	Low sensitivity	Cross-sectional difference
(1)	(2)	(3)	(4)	(5) = (2) - (4)
1 (Low)	-0.235	-0.071	0.097	
2	-0.308	-0.106	0.198	
3 (Moderate)	-0.277	-0.107	0.174	
4	-0.238	-0.043	0.248	
5 (High)	-0.144	0.002	0.121	
Difference (Low - Moderate)	0.042	0.036	-0.077	0.119
t-statistics	(1.04)	(0.74)	(-1.33)	(2.53)
Difference (High - Moderate)	0.133	0.109	-0.052	0.185
t-statistics	(2.94)	(2.05)	(-0.92)	(3.98)

Panel C: Illiquid stocks

Quintiles of sentiment	Average monthly synchronicity High sensitivity	Medium	Low sensitivity	Cross-sectional difference
(1)	(2)	(3)	(4)	(5) = (2) - (4)
1 (Low)	-0.345	-0.265	-0.095	
2	-0.368	-0.334	-0.041	
3 (Moderate)	-0.375	-0.308	-0.042	
4	-0.342	-0.288	0.016	
5 (High)	-0.256	-0.235	-0.104	
Difference (Low - Moderate)	0.030	0.043	-0.053	0.083
t-statistics	(1.07)	(1.26)	(-0.81)	(1.47)
Difference (High - Moderate)	0.119	0.073	-0.062	0.180
t-statistics	(3.60)	(1.86)	(-0.95)	(3.28)

Table 6
Cross-sectional variation in the sentiment-synchronicity relation: Multiple regressions

This table reports the results from estimating regressions of the following form:

$$Synchronicity_{j,t} = c_0 + c_1 SentResPos_{t-1} + c_2 AbsSentResNeg_{t-1} + c_3 SentResPos_{t-1} \times Rank6V_{j,t-1} + c_4 GDP_{t-1}$$
$$+ c_5 GDP_{t-1} \times Rank6V_{j,t-1} + c_6 Illiq_{t-1} + c_7 Illiq_{t-1} \times Rank6V_{j,t-1} + c_8 PeriodCount_t + c_9 PeriodCountSq_t$$
$$+ c_{10} ROESynchronicity_{j,t-1} + \varepsilon_{j,t}$$

The subscript *j* in the equation indexes firms, whereas subscript *t* indexes months. The equation is estimated using pooled, cross-sectional and time-series data. The dependent variable, *FF4F_Synchronicity*, is based on the log transformation of the ratio of R-squared to one minus R-squared of a regression in which stock returns are regressed on the Fama-French three factors and the momentum factor. *SentRes* represents the Baker-Wurgler sentiment index that has been orthogonalised to the T-bill rate and Lee's (2011) liquidity factor. *SentResPos* takes the value of *SentRes* when it is positive and zero otherwise. *AbsSentResNeg* takes the absolute value of *SentRes* when it is negative and zero otherwise. *Rank6V* is the composite rank based on the first principal component of the ranks of all stocks based on six characteristics: size, age, stock return volatility, dividend payout, MB ratio and the stock price level. Low values of *Rank6V* signify greater sensitivity to investor sentiment. *GDP* is the growth rate in real gross domestic product. *PeriodCount* and *PeriodCountSq* represent linear and quadratic time effects, respectively. *ROE Synchronicity* is the synchronicity of quarterly ROE, estimated from the most recent 12 quarters. *PctZero* is the per cent of zero-return days in the month. *Illiq* is the Amihud's (2002) measure of illiquidity. All results are obtained using monthly observations from October 1966 to January 2011. The *t*-statistics are reported in parentheses and are computed based on standard errors clustered by both firm and time (Thompson 2011).

	Dependent variable = *FF4F_Synchronicity*		
Explanatory variables	(1)	(2)	(3)
Intercept	-0.019	0.015	0.002
	(-0.28)	(0.22)	(0.03)
SentResPos	0.103	0.137	0.108
	(3.74)	(4.75)	(3.84)
AbsSentResNeg	0.002	0.007	-0.003
	(0.06)	(0.19)	(-0.08)
SentResPos x Rank6V	-0.010	-0.017	-0.010
	(-4.56)	(-6.65)	(-4.65)
PctZero (Firm-level, cross-sectionally demeaned)		0.001	
		(3.44)	
PctZero (Firm-level, cross-sectionally demeaned) x Rank6V		-0.001	
		(-23.76)	
Illiq (Firm-level, cross-sectionally demeaned)			0.0003
			(1.23)
Illiq (Firm-level, cross-sectionally demeaned) x Rank6V			-0.001
			(-9.37)
GDP	0.021	0.029	0.019
	(1.92)	(2.51)	(1.74)
*GDP*Rank6V*	-0.011	-0.013	-0.010
	(-6.95)	(-6.97)	(-6.38)
PeriodCount	-0.005	-0.005	-0.005
	(-11.13)	(-11.54)	(-11.26)
PeriodCountSqr	0.000	0.000	0.000
	(12.72)	(13.30)	(12.90)
Rank6V	0.071	0.061	0.065
	(30.75)	(24.40)	(28.91)
ROE Synchronicty	0.006	0.005	0.005
	(4.87)	(4.69)	(4.37)
Adj R^2	0.096	0.110	0.100
Number of pooled stock-month obs.	1,733,113	1,733,094	1,714,140

Table 7

Robustness checks: *DOX* and market volatility as additional controls

This table reports robustness checks for the results in Table 3 for aggregate market synchronicity. We add additional controls – Cassella and Gulen's (2018) index of extrapolative beliefs (*DOX*) and market volatility (*MktVol*) – to multiple regressions of aggregate stock return synchronicity on investor sentiment. The dependent variable, *FF4F_Synchronicity*, is based on the log transformation of the ratio of R-squared to one minus R-squared of a regression in which stock returns are regressed on the Fama-French three factors and the momentum factor. *SentPos* takes the value of the Baker-Wurgler sentiment index (*Sent*) when it is positive and zero otherwise. *AbsSentNeg* takes the absolute value of *Sent* when it is negative and zero otherwise. *SentResPos* takes the value of the orthogonalised Baker-Wurgler sentiment index (*SentRes*) when it is positive and zero otherwise. *AbsSentResPos* takes the absolute value of *SentRes* when it is negative and zero otherwise. *SentRes* represents the Baker-Wurgler sentiment index that has been orthogonalised with respect to the T-bill rate and Lee's (2011) liquidity factor. *GDP* is the growth rate in real gross domestic product. *PeriodCount* and *PeriodCountSq* represent linear and quadratic time effects, respectively. *ROE Synchronicity* is the synchronicity of quarterly ROE, estimated from the most recent 12 quarters. *D/P* is the aggregate dividend-price ratio. All results are obtained using monthly observations. The sample period is from October 1966 to January 2011. The *t*-statistics are reported in parentheses and are based on Newey-West (1987) standard errors.

	Dependent variable			
	FF4F_Synchronicity			*FF4F_ Synchronicity (MV-weighted)*
Explanatory variables	(1)	(2)	(3)	(4)
Intercept	0.048	0.052	0.025	0.186
	(0.68)	(0.71)	(0.45)	(1.76)
DOX	0.242	0.184		0.286
	(2.62)	(1.74)		(1.72)
D/P	-4.422	-9.172		24.225
	(-0.35)	(-0.68)		(1.56)
*DOX*D/P*	1.395	13.360		-29.447
	(0.05)	(0.44)		(-0.84)
SentResPos		0.057	0.076	-0.030
		(2.03)	(3.20)	(-0.79)
AbsSentResNeg		0.008	0.025	0.068
		(0.18)	(0.62)	(1.17)
MktVol (%)			0.007	
			(6.85)	
GDP	-0.014	-0.013	-0.008	-0.044
	(-1.16)	(-1.13)	(-0.85)	(-2.36)
PeriodCount	-0.003	-0.003	-0.003	-0.001
	(-7.07)	(-6.38)	(-7.23)	(-1.03)
PeriodCountSqr	0.000005	0.000006	0.000006	0.000002
	(7.03)	(6.56)	(7.88)	(1.86)
ROE Synchronicity	0.028	0.014	0.010	-0.021
	(1.00)	(0.46)	(0.39)	(-0.59)
Adj R^2	0.27	0.33	0.35	0.18
No. of months	517	517	517	517

Table 8

Robustness checks: Panel regressions using alternative sentiment indexes, exclusion of the bubble period and contemporaneous relationship

This table reports the robustness checks on the panel regression analysis reported in Table 6. As in Table 6, we regress stock-level return synchronicity on investor sentiment and control variables. The dependent variable, *FF4F_Synchronicity*, is based on the log transformation of the ratio of R-squared to one minus R-squared of a regression in which stock returns are regressed on the Fama-French three factors and the momentum factor. The robustness checks in specifications (1) and (2) use alternative sentiment indexes. In specification (1), we construct an index following the methodology of Baker and Wurgler (2006, 2007), but apply it recursively (rather than on the full sample) to avoid look-ahead biases. This index is available from December 1984 to December 2010. We orthogonalise this index to the T-bill rate and Lee's (2011) liquidity factor. *RecSentResPos* takes the value of the orthogonalised recursive index when it is positive and zero otherwise. *AbsRecSentResNeg* takes the absolute value of the orthogonalised recursive index when it is negative and zero otherwise. In specification (2), we use the PLS-based index constructed by Huang, Jiang, Tu and Zhou (2015). *SentPos_PLS* takes the value of the PLS-based index when it is positive and zero otherwise. *AbsSentNeg_PLS* takes the absolute value of the index when it is negative and zero otherwise. In specification (3), the Internet bubble period of January 1997 to March 2000 has been excluded. Specification (4) examines the contemporaneous relationship between synchronicity and sentiment. *SentRes* represents the Baker-Wurgler sentiment index that has been orthogonalised to the T-bill rate and Lee's (2011) liquidity factor. *SentResPos* takes the value of *SentRes* when it is positive and zero otherwise. *AbsSentResNeg* takes the absolute value of *SentRes* when it is negative and zero otherwise. *Rank6V* is the composite rank based on the first principal component of the ranks of all stocks based on six characteristics: size, age, stock return volatility, dividend payout, MB ratio and the stock price level. Low values of *Rank6V* signify greater sensitivity to investor sentiment. *GDP* is the growth rate in real gross domestic product. *PeriodCount* and *PeriodCountSq* represent linear and quadratic time effects, respectively. *ROE Synchronicity* is the synchronicity of quarterly ROE, estimated from the most recent 12 quarters. All results are obtained using monthly observations from October 1966 to January 2011. The *t*-statistics are reported in parentheses and are computed based on standard errors clustered by both firm and time (Thompson 2011).

	Dependent variable = *FF4F_Synchronicity*			
	Alternative sentiment indexes		Excluding Internet bubble period	Contemporaneous relationship
Explanatory variables	(1)	(2)	(3)	(4)
Intercept	-0.562	-0.007	-0.035	-0.041
	(-13.75)	(-0.15)	(-0.49)	(-0.59)
RecSentResPos	0.084			
	(2.14)			
AbsRecSentResNeg	-0.038			
	(-0.74)			
RecSentResPos x Rank6V	-0.014			
	(-4.53)			
SentPos_PLS		0.043		
		(2.64)		
AbsSentNeg_PLS		-0.044		
		(-1.81)		
SentPos_PLS x Rank6V		-0.004		
		(-2.40)		
SentResPos			0.126	0.097
			(4.30)	(3.46)
AbsSentResNeg			0.036	0.017
			(1.01)	(0.41)
SentResPos x Rank6V			-0.011	-0.011
			(-4.96)	(-5.15)
GDP	0.055	0.017	0.018	0.019
	(3.11)	(1.57)	(1.60)	(1.73)
*GDP*Rank6V*	-0.018	-0.011	-0.01	-0.011
	(-7.66)	(-6.88)	(-6.32)	(-6.85)
PeriodCount	-0.003	-0.005	-0.005	-0.005
	(-5.72)	(-13.15)	(-10.17)	(-10.45)
PeriodCountSqr	0.00002	0.00001	0.00001	0.00001
	(9.51)	(14.57)	(11.63)	(11.99)
Rank6V	0.080	0.070	0.073	0.072
	(28.01)	(33.25)	(30.03)	(30.93)
ROE Synchronicty	0.005	0.006	0.006	0.006
	(4.12)	(4.87)	(4.10)	(4.97)
Adj R^2	0.118	0.096	0.098	0.095
Number of pooled stock-month obs.	1,377,590	1,733,113	1,519,178	1,729,724

Appendix A
Description of the variables

Variables (notation)	Unit of computation	Definition
		Panel A: Measures of stock return synchronicity
MYY measure of stock return synchronicity (*MYY_Synchronicity*)	Stock-month	Measure of stock return synchronicity developed by Morck, Yeung and Yu (2000). It is computed for each stock-month, by regressing the daily stock returns on contemporaneous and lagged market and industry returns. Industries are defined according to the Fama-French 17-industries classification scheme. Synchronicity is then obtained by taking the natural log of the ratio between R^2 and 1-$R^2$of the regression, as specified in Equation (2). Stock-level synchronicity is aggregated into market-level synchronicity using Equations (3) and (4). We remove stock-month observations for which the stock has missing return data in more than 10% of all trading days during the quarter.
Fama-French-Carhart four-factor stock return synchronicity (*FF4F_Synchronicity*)	Stock-month	Computed in the same fashion as *MYY_Synchroncity* except that the return model used is the Fama-French-Carhart four-factor modol, which includes the three Fama-French factors as well as the momentum factor.
Fama-French five-factor stock return synchronicity (*FF5F_Synchronicity*)	Stock-month	Computed in the same fashion as *MYY_Synchroncity* except that the return model used is the Fama-French five-factor model, which includes the three Fama-French factors as well as the profitability and investment factors.
Fama-French global four-factor stock return synchronicity (*FF4GF_Synchronicity*)	Stock-month	Computed in the same fashion as *MYY_Synchroncity* except that the return model employs the global versions of the Fama-French three factors and the momentum factor. As these data are available only from 1991, we use the global factors from January 1991 onwards and the local US factors for the earlier period.

Appendix A (continued)

Variables (Notation)	Unit of computation	Definition
		Panel B: Measures of investor sentiment
Investor sentiment index (*Sent*)	month	Index of investor sentiment created by Baker and Wurgler (2007) and downloaded from http://pages.stern.nyu.edu/~jwurgler/. It is based on the first principal component of six standardised sentiment proxies, in which each of the proxies has first been orthogonalised with respect to a set of macroeconomic conditions. The index provided by Baker and Wurgler is available monthly. We match the value of the index for month *t* with synchronicity in month *t*+1.
Residual investor sentiment index (*SentRes*)	month	We follow the procedure in Sibley, Wang, Xing and Zhang (2016) and regress the Baker-Wurgler monthly sentiment index on the T-bill rate and Lee (2011) liquidity risk factor. The residuals of the regression represent the residual measure of investor sentiment.
PLS invetsor sentiment index (*Sent_PLS*)	month	Index of investor sentiment created by Huang, Jiang, Tu and Zhou (2015) and downloaded from http://apps.olin.wustl.edu/faculty/zhou/. Hung, Jiang, Tu and Zhou (2015) use PLS methodology to extract the index from the same six proxies used by Baker and Wurgler.
Degree of extrapolation bias in investor beliefs (*DOX*)	month	Index developed by Cassella and Gulen (2018) using the Investor Intelligence Survey (II) and the American Association of Individual Investors (AA). The index depicts the degree of extrapolative weighting in investors' belief of future returns. It is estimated recursively using principal components of the II and AA surveys. We extract approximate values of it from Figure 4A of Cassella and Gulen (2018) with the help of a digitising software and linear interpolation.
		Panel C: Control variables
Real GDP growth in per cent (*GDP*)	month	The first difference of the logarithm of quarterly real GDP, obtained from the US Bureau of Economic Analysis.

Appendix A (continued)

Variables (Notation)	Unit of computation	Definition
Synchronicity of *ROE* *(ROE_Synchronicity)*	Firm-month	At each quarter-end, we compute the synchronicity of firm-level ROE with industry- and market-level ROE, using data from the 12 most recent quarters. ROE is computed as Income Before Extraordinary Items (IBQ) divided by Total Shareholders' Equity (SEQQ), with both variables being obtained from Compustat. In each quarter we classify firms into 17 industries based on the industry classification of Fama and French. We then compute value-weighted market and industry ROE for each quarter. We regress ROE of a firm on contemporaneous value-weighted market and industry ROE and compute the R^2. ROE synchronicity is then obtained by taking the natural log of the ratio between R^2 and 1-R^2. To be included in the analysis, firms are required to have at least 4 quarterly ROE observations over the 12 most recent quarters.
3-month T-bill rate (*Tbill*)	month	T-bill rate is the 3-month Treasury bill rate. The data were downloaded from the WRDS.
Per cent of zero return (*PctZero*)	Stock-month	This is the per cent of zero-return days for the stock in a month. We use daily return data from CRSP to compute this measure. We also aggregate the firm-level data to obtain aggregated market-level *PctZero* by taking a simple average.
Amihud illiquidity (*Illiq*)	Stock-month	Amihud (2002) measure of illiquidity is computed using daily returns and volume data from CRSP. For each day, we compute the ratio of absolute return to dollar trading volume and then sum this ratio over the month to obtain monthly values of Amihud Illiquidity for the stock.
Market volatility (*MktVol*)	month	Annualised standard deviation of market daily return (obtained from CRSP) within each month.
Aggregate dividend to price ratio (*D/P*)	month	Aggregated dividend to price ratio, obtained from CRSP.

Panel D: Cross-sectional rank variables

Rank of firm size (*Ranksize*)	Stock-month	Month-end rank of a firm's market capitalisation (1-smallest firm, 10-largest firm), using NYSE breakpoints.
Rank of firm age (*Rankage*)	Stock-month	Month-end rank of a firm's age, defined as the number of months since its inclusion in the CRSP database (1-youngest, 10-oldest). Breakpoints for ranks are identified using NYSE stocks.

Appendix A (continued)

Variables (Notation)	Unit of computation	Definition
Rank of stock return volatility (*RankStd*)	Stock-month	Month-end rank of a stock's return volatility, defined as the standard deviation of its monthly returns over the most recent 12-month period (1-most volatile, 10-least volatile). We drop stocks that do not have at least 9 monthly return observations. Breakpoints for ranks are identified using NYSE stocks.
Rank of dividend-to-book ratio (*RankDiv*)	Stock-month	Annual rank of a firm based on its dividend-to-book ratio (0-non-payers, 1-lowest non-zero ratio, 10-highest ratio), as of the most recent June end. Dividends are defined as the product of dividends per share at the ex date (Compustat item DVPSXQ) and number of common shares outstanding (Compustat item CSHOQ). We exclude observations with negative book value of equity. Breakpoints for ranks are identified using NYSE stocks.
Market-to-book ratio (*RankMB*)	Stock-month	Annual rank based on a firm's market-to-book ratio (MB) as at the most recent June end. MB is defined as market capitalisation (Compustat items PRCC_C * CSHO) divided by Total Shareholders' Equity (Compustat item SEQ). We exclude observations with zero or negative book value of equity. Breakpoints for ranks are identified using NYSE stocks. We first sort all stocks into deciles based on their MB ratio. We then assign a rank of 1.5 to stocks in the top and bottom deciles, a rank of 3.5 to stocks in deciles 2 and 9, and so on. Finally, stocks in the middle two deciles are assigned a rank of 9.5. This ranking scheme allows us to test whether the effects of sentiment increase when MB becomes more extreme.
Rank of stock price (*RankPrc*)	Stock-month	Month-end rank of a firm's stock price level (1-lowest price, 10-highest price). Breakpoints for ranks are identified using NYSE stocks.
Composite rank of the six rank variables (*Rank6V*)	Stock-month	Following Mian and Sankaraguruswamy (2012), the composite rank is based on the first principle component of the decile ranks of the six individual characteristics: size, age, stock return volatility, dividend-to-book ratio, MB ratio and stock price. The methodology of computing the ranks on the six individual characteristics is described above. Lower composite rank values indicate greater sensitivity to investor sentiment.

Appendix B

A model of aggregate investor sentiment and stock return synchronicity

We use a modified version of Daniel, Hirshleifer and Subrahmanyam's (1998) model to show how aggregate investor sentiment affects stock return synchronicity. We consider a risky stock with three dates: $t = 0, 1, 2$. The stock pays a liquidating dividend, *D*, at t = 2, when consumption occurs. D is not known at $t < 2$, is assumed to be lognormally distributed, and is dependent on an aggregate factor, m, and an idiosyncratic factor, *v*:

$$logD = d = \beta m + v, \quad \text{(B.1)}$$

where β is the stock's loading on the aggregate factor. Both *m* and *v* are normally distributed and independent from each other. The representative investor is risk neutral and can borrow and lend at a risk-free, 0% interest rate. The investor's prior beliefs on *m* and *v* at date 0 are given by:

$$m \sim N(0, \sigma_m^2), v \sim N(0, \sigma_v^2).$$

On date 1, the investor observes two signals about the two factors:

$$s_{1,m} = m + \epsilon_m, s_{1,v} = v + \epsilon_v,$$

where ϵ_m and ϵ_v are random noise in the two signals. ϵ_m, ϵ_v, *m* and *v* are all independent from each other. ϵ_m and ϵ_v are normally distributed as:

$$\epsilon_m \sim N\left(0, \sigma_{1,m}^2\right), \epsilon_v \sim N\left(0, \sigma_{1,v}^2\right).$$

We assume that aggregate investor sentiment affects the investor's learning process on date 1 with respect to the aggregate factor, *m*. In particular, when aggregate sentiment is high, the investor overestimates the precision of the aggregate signal $s_{1,m}$.

B.1 Normal sentiment period

During normal sentiment periods, the investor correctly perceives the precision of both signals. Using Bayes' rule to update her prior, the investor's posterior beliefs at date 1 about *m* and *v* are normally distributed, with means given by:

$$E\left(m \middle| s_{1,m}\right) = \frac{\frac{1}{\sigma_{1,m}^2}}{\frac{1}{\sigma_m^2} + \frac{1}{\sigma_{1,m}^2}} s_{1,m}, \text{ and } E\left(v \middle| s_{1,v}\right) = \frac{\frac{1}{\sigma_{1,v}^2}}{\frac{1}{\sigma_v^2} + \frac{1}{\sigma_{1,v}^2}} s_{1,v}.$$

Because the investor is risk neutral, the stock price is determined by the investor's expected payoff on each date. Denoting the log stock price on each date as p_t, it can be shown that $p_0 = 0$,

$$p_1 = \beta E\left(m \middle| s_{1,m}\right) + E\left(v \middle| s_{1,v}\right) = \beta \frac{\frac{1}{\sigma_{1,m}^2}}{\frac{1}{\sigma_m^2} + \frac{1}{\sigma_{1,m}^2}} s_{1,m} + \frac{\frac{1}{\sigma_{1,v}^2}}{\frac{1}{\sigma_v^2} + \frac{1}{\sigma_{1,v}^2}} s_{1,v}, \qquad \text{and} \qquad p_2 = \beta m + v.$$

Because all shocks are normally distributed and additive, the liquidating dividend and all prices are lognormally distributed. It follows that log price changes (Δp) are returns, and the volatility of log price changes are return volatilities. Over the full horizon from $t = 0$ to $t = 2$, the stock's synchronicity, or its R^2 with the aggregate factor, is given by

$$R^2 = \frac{\beta^2 \sigma_m^2}{\beta^2 \sigma_m^2 + \sigma_v^2}. \tag{B.2}$$

B.2 High sentiment period

We postulate that aggregate investor sentiment affects investor learning. Our goal is to capture the effects of sentiment that is of a market-wide nature, rather than specific to certain segments of the market. When aggregate sentiment is high, the investor overestimates the precision of the aggregate signal $s_{1,m}$ by a multiplicative factor of $\rho > 1$. At t = 1, the investor still uses Bayes' rule to update her prior, but her posterior beliefs are now given by:

$$E(m|s_{1,m}) = \frac{\rho \frac{1}{\sigma_{1,m}^2}}{\frac{1}{\sigma_m^2} + \rho \frac{1}{\sigma_{1,m}^2}} s_{1,m}, and\ E(v|s_{1,v}) = \frac{\frac{1}{\sigma_{1,v}^2}}{\frac{1}{\sigma_v^2} + \frac{1}{\sigma_{1,v}^2}} s_{1,v}.$$

This result suggests that aggregate sentiment leads the investor to overreact to the aggregate signal. In fact, the higher the bias, ρ, the stronger the investor's reaction coefficient $\frac{\rho \frac{1}{\sigma_{1,m}^2}}{\frac{1}{\sigma_m^2} + \rho \frac{1}{\sigma_{1,m}^2}}$ to the aggregate signal, $s_{1,m}$. Such reactions by the investor lead to the following stock prices across the three dates: $p_0 = 0, p_1 = \beta E(m|s_{1,m}) + E(v|s_{1,v}) = \beta \frac{\rho \frac{1}{\sigma_{1,m}^2}}{\frac{1}{\sigma_m^2} + \rho \frac{1}{\sigma_{1,m}^2}} s_{1,m} + \frac{\frac{1}{\sigma_{1,v}^2}}{\frac{1}{\sigma_v^2} + \frac{1}{\sigma_{1,v}^2}} s_{1,v}$, and $p_2 = \beta m + v$. These stock price dynamics imply that, over the full horizon from t = 0 to t = 2, the aggregate component of total return variance is given by $\sigma_{2,m}^2 = \beta^2 \sigma_m^2 + \frac{2\beta^2 \rho(\rho - 1) \frac{1}{\sigma_{1,m}^2}}{(\frac{1}{\sigma_m^2} + \rho \frac{1}{\sigma_{1,m}^2})^2}$ and the firm-specific component of total return variance is still given by σ_v^2. The stock's R^2 with the aggregate market factor is then given by:

$$R^2 = \frac{\sigma_{2,m}^2}{\sigma_{2,m}^2 + \sigma_v^2}, \tag{B.3}$$

$$\text{where } \sigma_{2,m}^2 = \beta^2\sigma_m^2 + \frac{2\beta^2\rho(\rho-1)\frac{1}{\sigma_{1,m}^2}}{(\frac{1}{\sigma_m^2}+\rho\frac{1}{\sigma_{1,m}^2})^2}.$$

When $\rho > 1$, market comovement increases both when we go from t = 0 to t = 1 (as the investor overreacts to the aggregate signal at t = 1) and when we go from t = 1 to t = 2 (as uncertainty is resolved). In fact, it can be shown that if the investor overestimates the precision of the aggregate signal, $s_{1,m}$, by a multiplicative factor of $\rho > 1$, the stock's return synchronicity, or its R^2 with the aggregate factor, is increasing in ρ.

The time-series implication of this result is that stock return synchronicity tends to increase when market sentiment strengthens and the bias coefficient ρ is high. But this result can also have cross-sectional implications. Baker and Wurgler (2006, 2007) suggest that the effect of investor sentiment can vary across stocks. Stocks that are more difficult to value and arbitrage tend to be more severely affected by market sentiment as it is more costly and risky for rational, non-sentimental arbitrageurs to trade against sentimental investors in these stocks. As a result, at the same point in time, difficult to arbitrage stocks will tend to have a higher bias coefficient, ρ, than easy-to-arbitrage stocks. To test for the presence of such cross-sectional differences, we examine if the effect of aggregate market sentiment on stock return synchronicity is more pronounced among companies that are small, young, volatile, non-dividend paying, low-priced and with extreme valuation ratios – firms that Baker and Wurgler (2006, 2007) suggest have more subjective valuations and are more difficult to arbitrage.

Finally, although the dividend process (B.2) contains only a single factor (the market factor), this assumption is made without loss of generality. Even if there are in fact multiple factors in the dividend process, as long as investors' learning with respect to these factors remains analogous to that for the market factor, our conclusion will still be valid.[18] Specifically, if there are factor-specific signals that investors learn over time and investors overestimate their precision during high sentiment periods, investors will overreact to the factor signals during these times. Such overreaction and subsequent reversal (after uncertainty is resolved), in turn, generates excess comovement and higher R^2 with the (multiple) aggregate factors. Furthermore, if the degree of sentiment-induced overreactions are disproportionately stronger for certain stocks (e.g. difficult to arbitrage stocks), their increase in synchronicity during high sentiment periods will also be more substantial.

[18] Ball, Sadka and Sadka (2009) show that firms' earnings fundamentals and equity returns follow similar factor structure by using a principal-components analysis.