Analyst herding-whether, why, and when? Two new tests for herding detection in target forecast prices.

REVELEY, C., SHANAEV, S., BIN, Y., PANTA, H. and GHIMIRE, B.

2023

sciendo
Economics and Business Review
https://www.ebr.edu.pl
Vol. 9 (4), 2023: 25–55
https://doi.org/10.18559/ebr.2023.4.892

Analyst herding—whether, why, and when? Two new tests for herding detection in target forecast prices

Callum Reveley[1] Savva Shanaev[2] Yu Bin[3]
Humnath Panta[4] Binam Ghimire[5]

Abstract

This study proposes two novel tests for security analyst herding based on binomial correlation and forecast error volatility scaling, and applies it to investigate herding patterns in analyst target prices in 2008–2020 in the UK. Analysts robustly herd in their valuations, with results consistent across years, sectors, in terms of panel fixed effect, quantile, instrumental variable regressions, and when controlled for optimism and conservatism. Herding becomes prominent for stocks followed by at least five analysts and towards the long sides of Fama-French sorts, reinforcing its non-spurious and behavioral nature.

Keywords

- behavioral finance
- financial econometrics
- herding
- stock analyst

JEL codes: C58, G23, G41

Article received 10 September 2023, accepted 12 December 2023.

Suggested citation: Reveley, C., Shanaev, S., Bin, Y., Panta, G., & Ghimire, B. (2023). Analyst herding—whether, why, and when? Two new tests for herding detection in target forecast prices. *Economics and Business Review*, *9*(4), 25–55. https://doi.org/10.18559/ebr.2023.4.892

[1] University of Northumbria at Newcastle, Newcastle-upon-Tyne, NE1 8ST, United Kingdom, callum2.reveley@northumbria.ac.uk.

[2] University of Northumbria at Newcastle, Newcastle-upon-Tyne, NE1 8ST, United Kingdom, s.shanaev@northumbria.ac.uk, https://orcid.org/0000-0002-8031-3490.

[3] Zhejiang University of Technology, Jianxing Building, 288 Liuhe Rd, Hangzhou, 310023, PR China, yubin@zjut.edu.cn, https://orcid.org/0000-0002-6870-099X.

[4] School of Business, SH 126, Cal Poly Humboldt, 1 Harpst St, Arcata, CA 95519, USA, Humnath.Panta@humboldt.edu, https://orcid.org/0000-0002-3460-2371.

[5] Robert Gordon University, Garthdee House, Garthdee Road, Aberdeen, AB10 7QB, Scotland, United Kingdom, corresponding author: b.ghimire@rgu.ac.uk, https://orcid.org/0000-0003-1351-7247.

Introduction

The question of whether, why, how and when security analysts herd in their forecasts and valuations has been a subject of active and intense academic debate at least since the early 1990s (Scharfstein & Stein, 1990). The literature on institutional and analyst herding has substantially expanded since then, with a plethora of theoretical and empirical studies having emerged (see, e.g., Clement & Tse, 2003, 2005; Frijns & Huynh, 2018; Hong & Kacperczyk, 2010; Lee & Lee, 2015; Trueman, 1994; Welch, 2000). However, researchers disagree on the causes and consequences of herding or sometimes with regard to its very existence. Early research either hypothesised that herding is purely irrational and of a behavioral nature (Welch, 2000), or reputational, stemming from information asymmetry and differing ability of analyst (Scharfstein & Stein, 1990; Clement & Tse, 2005). Further research proposed a conflict-of-interest explanation, introducing a principal-agent problem dimension into the analysis (Hong & Kacperczyk, 2010; James & Karceski, 2006; Lee & Lee, 2015), and analyst competition as the herding-mitigating factor. Other studies argue that herding is spurious due to analysts and market participants relying on similar valuation models and the same fundamental information or demonstrating other biases, yielding herding nothing more than a statistical artefact (Guo et al., 2020). Hence, the literature does not reach a consensus on the issue, with assessing the presence and the degree of herding being increasingly difficult both theoretically and econometrically, and often requiring large, specialised, disaggregated, high-frequency, and analyst-level datasets (Bernhardt et al., 2006; Blasco et al., 2018).

Therefore, this study seeks to contribute to the literature on analyst herding by developing two conceptually and computationally simple yet flexible and powerful econometric tests capable of determining herding patterns in analyst target forecast prices using non-specialised and aggregated data and introducing a battery of robustness tests generating testable implications for the competing theories of analyst herding, and applying them to the UK stock market in 2008–2020, utilising a sample of over 2,000 followed companies, over 12,000 stock-year observations, and in excess of 85,000 individual analyst forecasts. We propose non-parametric and parametric tests. The non-parametric test has its foundations in a binomial default correlation framework and the second, the parametric one, is the main test we propose here. This second test exploits the logic of variance scaling for dependent and independent variables.

This study establishes that analyst herding is robustly present in analyst valuations. The findings are consistent with the behavioral theories of herding rather than conflict-of-interest or reputational theories; they reinforce its non-spurious nature, and highlight the greater prominence of herding behaviour subject to lower volatility and uncertainty. The results persist in subsamples, and when concerns regarding heterogeneity, endogeneity, spurious herding, and outliers are addressed. The contribution of this study is therefore two--fold, highlighting both the financial econometrics of herding detection and the policy implications of the observed herding patterns.

The rest of the paper is organised as follows: First, the literature on herding in individual stock market participants, institutional investors, and security analysts is reviewed. Next, the data this study utilises is presented and the two econometric tests it applies are derived. The findings section presents the estimation results and robustness checks, discussing them in the context of the literature and the competing theories of analyst herding. The final section presents conclusions.

1. Literature review

The literature on individual investor, analyst, and institutional herding is yet to reach a consensus on whether herding or contrarianism are more prominent on financial markets. An agreement is also missing on their implications for market efficiency and quality, and with regards to the robustness of herding, its spurious, reputational, or behavioral nature, and interactions with other behavioral biases such as overconfidence, optimism, and conservatism. This state of affairs is confirmed by generally mixed and inconclusive findings in both econometric and experimental studies.

Early research on analyst target forecasts tended to illuminate their informational value and implications for investing. O'Brien (1988) showed that, on average, analyst earnings predictions are better than those of econometric time-series models. However, there was also evidence of analyst conservatism, with lagged analyst forecast errors having predictive power over current earnings. Doukas et al. (2005) report more nuanced findings: while analyst coverage in their sample alleviates agency problems, it also leads to persistent over- and under-valuations of strongly and weakly covered stocks, respectively. Imam et al. (2013) show that analyst target price forecast accuracy depends on the underlying valuation model, with those forecasts based on return on equity and book value performing the best. Brav and Lehavy (2003) document the informativeness of analyst target forecast prices by studying market reactions to forecast announcements and revisions, while also reporting that one-year-ahead target prices are overly optimistic. While numeric information such as earnings management related can have an impact on valuation (Kałdoński & Jewartowski, 2017), Huang et al. (2009) demonstrate that analyst recommendations contain substantial information not available in numerical valuations, with strategies synthesising forecasts and recommendations delivering higher returns. Similarly, Feldman et al. (2012) argue that the best-performing strategy should incorporate both target price and earnings forecasts, as well as analyst recommendations. Ishigami and Takeda (2018) state that the accuracy of forecasts is conditional on the quality of security analysis firms. Moreover, Han et al. (2021) develop a robust measure of market reaction to target price announcements, showing a short-term price appreciation and a subsequent reversal. Recent literature has become more critical of analyst forecasting ability, with Bradshaw, Brown et al. (2013) and Bradshaw, Huang et al. (2014) showing how analysts are persistently optimistic, with buy-side analysts' ability slightly lower than that of sell-side analysts. Lin et al. (2016) establish a link between institutional trading and analyst recommendations, therein showing that trading activity increases subject to a target price revision, albeit such trades do not generate abnormal returns.

The concept of institutional herding has been introduced by Scharfstein and Stein (1990) in a theoretical model where managers can mimic the decisions of others, resulting in a privately optimal but socially suboptimal equilibrium. Scharfstein and Stein (1990) were inspired by both behavioral economics insights on group psychology and the corporate finance concept of the principal-agent problem, suggesting a reputational incentive to herd. In a seminal empirical paper, Trueman (1994) utilises a non-parametric test applied to sequential security analyst forecast releases to document herding in their valuations. Trueman (1994) shows that herding affects the apparent stock price reaction to earnings surprises and can lead to mismeasurements in conventional information dissemination models. Welch (2000) argues that herding leads to underestimation of volatility and increased fragility of financial markets and, as the degree of herding is not affected by the accuracy of prior consensus forecasts, the findings favour behavioral or informational asymmetry theories of the herding. Drehmann et al. (2005) design a laboratory experiment to study herding patterns and find that not overly consensual but rather contrarian valuations are more detrimental to market equilibrium stability, subject to informational cascades, while herding can be combatted with market design, particularly flexible pricing quotes. Additionally, they report herding patterns to be similar across market participants with varying roles. Contrastingly, in an empirical study, Clement and Tse (2005) argue that bold contrarian forecasts are more accurate and can be socially beneficial, while also relating herding and contrarianism to past analyst experience, ability, specialisation, and reputational incentives. Graham (1999) supports the reputational herding theory, showing that analysts tend to imitate their peers when their own reputation is not strong or when their private signal contradicts a strong public informational signal.

Roider and Voskort (2016) introduce employers into their experimental setting and confirm the reputational incentives for herding when rewards depend on forecast accuracy and ability is unobservable. Cote and Goodstein (1999) acknowledge the incentives which security analysts might have to herd and emphasise the ethical implications of herding and the virtue of bold contrarian forecasts. Chen and Jiang (2006) argue analysts are overconfident and contrarian in the sense they over-rely on their private information in comparison to private information. However, they show this is symptomatic of an incentive failure rather than persistent behavioral biases. Cheng et al. (2019) provide additional evidence in favour of agency-based analyst herding explanations, showing that target price accuracy is higher for firms with better corporate governance practices. Frijns and Huynh (2018) further confirm the privately rational nature of herding by exploiting the differential impact media sentiment has on analyst recommendations conditional on individual analyst characteristics. While not directly linked with the role of the analyst, herding is also found to be highly correlated with market sentiment, e.g., when it is measured by the VIX volatility index (Aharon, 2021).

Another strand in the literature investigates the effects of competition and conflict of interest including the role of volatility on analyst forecasts. James and Karceski (2006) provide evidence on excessively optimistic analyst coverage subject to underperforming IPOs, suggesting collusion between some analysts and the firm's underwriter. Hong and Kacperczyk (2010) and Wang et al. (2020) use broker mergers as natural experiments varying the level of competition between analysts, and document an increase in herding subject to such mergers. These findings are more consistent with behavioral than reputational herding explanations. Lee and Lee (2015) propose an explanation combining conflict of interest and informational asymmetry: analysts affiliated with the target stock are excessively optimistic and overly consensual, while others follow this signal as affiliated analysts can possess insider information. Loang and Ahmad (2021) emphasise the mediating role of volatility. They find that the release of analysts recommendation causes realised volatility to fluctuate and that investors are triggered by the volatility. They make use of realised volatility and the Parkinson estimator to measure it. Additionally, the literature tends to disagree on the impact that market conditions and other external variables have on the degree of herding between market participants. The conventional models developed by Christie and Huang (1995) and Chang et al. (2000) implicitly assume herding is most prominent subject to extreme market conditions. This is confirmed in early empirical research (Caparrelli et al., 2004) as well as in more recent analyst-focused studies, with Lin (2018) showing that herding intensifies when aggregate market uncertainty is high. Such uncertainties can be the result of several factors, including inefficiencies caused by the lack of experienced market participants or institutional rigidities restricting the flow of information (Wheeler et al., 2002). However, Hwang and Salmon (2004) find the opposite to be generally true, with herding associated to a greater extent with calm market periods. Welch (2000) also documents analyst herding towards the consensus being stronger under favourable market conditions. Galariotis et al. (2015) show herding is more pronounced when important macroeconomic or fundamental information is released, with this relationship heterogeneous across different markets.

Another major divide in the literature is manifested with regard to the herding of general stock market participants versus institutional investors and analysts. Since the 1990s, a wide range of powerful tests have been developed to detect herding of individual investors. These tests mainly make use of the stock price data, including the cross-sectional standard deviation (Christie & Huang, 1995), cross-sectional absolute deviation (Chang et al., 2000), and cross-sectional factor loading dispersion (Hwang & Salmon, 2004) tests, whose conceptual simplicity, wide applicability and flexibility has led to them gaining substantial popularity and enjoying continuous use in recent studies (see, e.g., Blake et al., 2017; Vidal-Tomas et al., 2019).

However, tests for the analyst and institutional herding suggested in the literature often require high frequency, specialised, or disaggregated data. As such, Welch (2000) studies the recency of real-time analyst forecast revisions to document herding. Bernhardt et al. (2006) propose a non-parametric S-statistic that conditions over- or undervaluations of individual analysts onto those of their peers. Friesen and Weller (2006) use Bayesian methods to assess consensus forecast precision by exploiting the ordering of analyst valuations. Nofsinger and Sias (1999), Sias (2004), Choi and Sias (2009), and Choi and Skiba (2015) exploit dynamic institutional holdings data to document herding in institutional investment decisions with regard to the US stock market, individual industries, and international financial markets. Guo et al. (2020) integrate individual analyst recommendations and institutional holdings to challenge prior findings and demonstrate that herding is most likely spurious.

Tests less demanding of the granularity of data, such as Olsen (1996) and De Bondt and Forbes (1999), which both exploit the shape of the forecast distribution and consensus estimate dispersion, have been subsequently criticised, as they are parametric and are thus not robust to cross-sectional correlations, irrational analyst optimism, and other behavioral biases (Blasco et al., 2018). This also corresponds to the findings of Hong and Kacperczyk (2010), who show that analyst forecasts in the absence of competition can be both excessively consensual and overly optimistic due to conflict of interest, and also to Nofsinger and Sias's findings (1999), who assert that the patterns in institutional holdings data are consistent with both herding and feedback trading by institutional investors. Additionally, Abarbanell and Lehavy (2003) argue that statistical artefacts in the distribution of analyst forecasts can explain most of the anomalies in the data that are usually interpreted in favour of analyst behavioral biases.

Therefore, there exists a notable gap in the financial econometrics literature on analyst herding, as no test so far has been developed that simultaneously a) can be applied to aggregated, low-frequency data; b) can distinguish between spurious, reputational, and behavioral herding; c) can be adjusted for other analyst behavioral biases; and d) addresses the conventional concerns surrounding parametric tests. This study seeks to address this gap by developing two tests for analyst herding—a non-parametric simplistic test inspired by binomial correlations, and a flexible regression-based test that is accommodative to a battery of robustness checks. The next section discusses the sample that this study utilises, and also provides the derivation and justification for the testing process

2. Data and methodology

2.1. The sample

This study considers an exhaustive sample of all stocks listed on the London Stock Exchange for at least one year in the time period 2008–2020 and that have had at least one security analyst issuing a target price forecast between 2008 and 2019. As analyst targets cover a 12-month, forward-looking period, target prices current as of 31st December are mapped onto closing stock prices as of 31st December the following year in order to calculate pricing errors and determine over- and undervaluation. Therefore, 2008–2019 analyst valuations correspond to the 2009–2020 market prices, respectively, with year-ends chosen to prevent overlapping of forecasts and to correctly associate price data with relevant annual fundamentals. The final sample constitutes 2,079 stocks, over 12,000 stock-year observations, and over 85,000 individual analyst valuations. All data used in this study were obtained from Bloomberg through the use of various relevant functions, such as “ANR” for analysts’ recommendations, which shows recommendations and predictions for selected stocks. The historical data is available on an aggregated basis only, implying that correlations between individual analyst forecasts are unobservable and must be estimated indirectly. However, each stock-year observation includes both the average target price and the number of analysts whose individual valuations were aggregated to obtain it, which is crucial for the estimation strategy of the parametric regression-based herding test developed further in this section.

Next, the stocks are further assigned their GICS sectoral classifications and annual Fama-French-style factor sorts across market beta, size, value, momentum, profitability, and investment. Market beta is measured daily against the FTSE 250 index, momentum is conventionally defined using 12-month prior returns, and the investment sort is executed based on annual asset growth, as in Fama and French (2015, 2018). Stocks are subsequently grouped into top 30%, middle 40%, or bottom 30% categories. Table 1 provides a snippet of the data as an example. The full raw data sample is available upon request from the corresponding author.

Table 1. Data example

Year	2016	2017	2018	2019	2020
Forecast	19.98	23.40	26.43	28.06	26.58
Price	22.43	24.80	23.08	22.35	12.98
Number of analysts	24	30	28	27	25
Sector	energy	energy	energy	energy	energy
Beta	high	mid	high	mid	mid
Size	large	large	large	large	large
Value	value	neutral	neutral	neutral	neutral
Momentum	loser	winner	sideways	sideways	sideways
Profitability	weak	mid	mid	mid	mid
Investment	mid	bottom	aggressive	conserva-tive	mid

Source: Bloomberg. Various functions available in Bloomberg were applied to obtain the data presented above, such as "ANR" for forecast, "PX_LAST" for share price etc.

Next, the estimation strategy used by the study for inference of herding from such aggregated stock analyst target forecasts data is presented.

2.2. The non-parametric test: Binomial correlations

The first test proposed by this study is a non-parametric approach. It builds upon the binomial default correlation framework first conceptualised and derived in the CreditMetrics framework (JPMorgan, 1997). CreditMetrics suggested a simple statistical technique to infer binomial default correlations between obligors within homogeneous default probabilities from the volatility of average default rate across subsamples:

$$\hat{\rho} = \frac{\sigma^2}{\mu - \mu^2} \sim T\left(0, \sqrt{\frac{1-\hat{\rho}^2}{n-2}}, n-2\right)$$

where $\hat{\rho}$ is the estimated binomial correlation, σ^2 is the variance of the proportion across subsamples, and μ is the full sample average. The statistical significance of $\hat{\rho}$ can then be assessed using a Student's T distribution with mean zero, standard deviation $\sqrt{\frac{1-\hat{\rho}^2}{n-2}}$, and degrees of freedom $n-2$, where n is the sample size. This study suggests applying this concept to correlations between analyst target prices (herding) instead, using a very natural extension of the method. While the correlation of analyst *valuations* is not directly observable from the aggregated data, the CreditMetrics procedure can allow to infer the correlation of analyst *overvaluations* – a binary variable equal to one if the average analyst target price is higher than the realised stock price 12 months forward, and zero otherwise. σ^2 here can be interpreted as the variance of the proportion of overvaluations across subsamples (years or sectors), and μ as the average sample proportion of overvaluations. As regards the binomial distribution properties, the results are equivalent if the proportions of undervaluations are considered instead. If $\hat{\rho}$ is statistically significant, the null hypothesis of independence can be rejected in favour of the alternative hypothesis of herding. Note that a positive binomial correlation implies a higher positive correlation between valuations, fully analogous to the CreditMetrics case relating asset return correlations to default correlations. This simple model is advantageous, as it is non-parametric and allows to estimate overvaluation correlations intuitively interpretable as the degree of analyst herding in a wide range of samples. The two notable shortcomings of the method are that it can only return positive correlations, by not allowing the alternative hypothesis of contrarianism to be tested, while also assuming the rate of overvaluation is homogenous across sample stocks. Therefore, by addressing these limitations and allowing for the implementation of more thorough robustness checks, this study also develops a regression-based parametric test for analyst herding, which is discussed further below.

2.3. The parametric test: Forecast error volatility scaling

The second and the main test proposed by the study exploits the parametric approach and the logic of variance scaling for independent and dependent variables. Consider the forecast variance for the aggregated average target price $\overline{P}_{it}^{A}$ A for stock i in year t produced by m analysts, whose valuations X_j are correlated and each have variance σ^2. For a correlation coefficient ρ between such valuations:

$$V(\overline{P}_{it}^{A}) = \sigma^2(m) = V\left(\frac{1}{m}\sum_{j=1}^{m} X_j\right) = \frac{1}{m^2}(m\sigma^2 + C_m^2\rho\sigma^2) = \frac{(1-\rho)\sigma^2}{m} + \rho\sigma^2$$

For independent valuations:

$$\rho = 0 \Rightarrow \sigma(\overline{P}_{it}^{A}) = \frac{\sigma}{\sqrt{m}} \Rightarrow ln\ ln\ \sigma(\overline{P}_{it}^{A}) = ln\ \sigma - 0.5\ ln\ m$$

This leads to the baseline regression estimation design:

$$ln\ ln \left| \frac{\overline{P}_{it}^{A}}{P_{it+1}} - 1 \right| = \alpha + \beta\ ln\ N_{it}^{A} + \varepsilon_{it}$$

where P_{it+1} is the one-year forward market price, N_{it}^{A} is the number of analysts covering stock i in year t (observed value of m), α is the natural logarithm of the individual analyst prediction error $ln\ \sigma$, β is the volatility scaling exponent, and ε_{it} is the error term. The null hypothesis of independence ($\rho = 0$, $\beta = -0.5$) can then be tested against two alternative hypotheses of herding ($\rho > 0$, $\beta > -0.5$) and contrarianism ($\rho < 0$, $\beta < -0.5$) via a T-test:

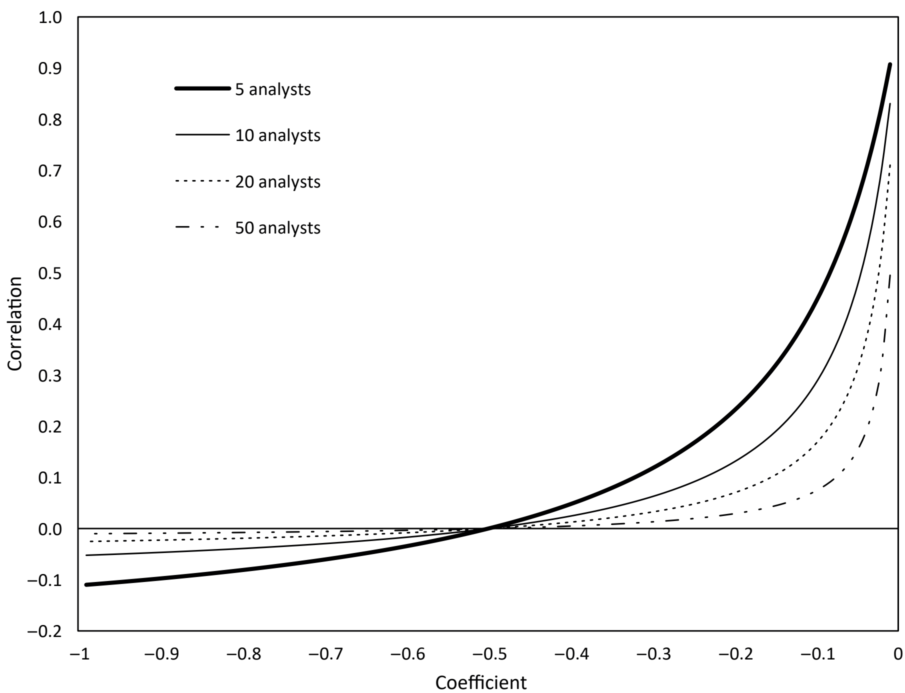

Figure 1. The correspondence between forecast volatility scaling and valuation correlations

Note: Plot created in MS Excel using data obtained from Bloomberg.

Source: Bloomberg.

$$t = \frac{\hat{\beta} + 0.5}{s(\hat{\beta})} \sim T(0, 1, n-2)$$

where $\hat{\beta}$ is the regression estimator of the volatility scaling exponent β, $s(\hat{\beta})$ is the heteroskedasticity-consistent standard error of $\hat{\beta}$ computed using the Huber-White covariance matrix, and n is the number of stock-year observations. The correspondence between the forecast volatility scaling estimator $\hat{\beta}$ and the valuation correlation coefficient $\hat{\rho}$ can be retrieved from the general-case relationship between $V(\overline{P}_{it}^{A})$ and N_{it}^{A}:

$$\beta = \frac{\partial \ln \sigma(\overline{P}_{it}^{A})}{\partial \ln N_{it}^{A}} = \frac{\rho - 1}{2(\rho N_{it}^{A} - \rho + 1)}$$

$$\hat{\rho} = -\frac{2\hat{\beta} + 1}{2\hat{\beta}(N_{it}^{A} - 1) - 1}$$

For the convenience of interpretation, the values of $\hat{\rho}$ for varying $\hat{\beta}$ coefficients and N_{it}^{A} are graphed in Figure 1. This allows to naturally construct confidence intervals for $\hat{\rho}$ alongside $\hat{\beta}$ to assess the magnitude and economic significance of herding or contrarianism between analysts more naturally and intuitively.

2.4. Robustness tests

The four major concerns that could compromise the validity of the results and thus are addressed in the robustness test employed by the study are heterogeneity, the impact of outliers, endogeneity, and spurious herding. Next, the procedures applied to address these are presented sequentially.

The heterogeneity of sample stocks is the major limitation for both the binomial correlation and the regression tests, as they assume the proportion of overvaluations μ and the individual analyst forecasting error σ, respectively, to be constant across the sample. If both individual analyst errors and analyst coverage are correlated with stock characteristics, the results obtained could be biased. To alleviate this concern, this study additionally conducts both tests in sectoral and yearly subsamples. For the parametric regression--based test, it also considers subsample estimations for Fama-French factor and panel regressions with the sector, year, and individual stock fixed effects. Such heterogeneity tests in subsamples are quite rare in the existing literature. The most notable findings here correspond to small stocks: Caparrelli et al. (2004) find herding in small stocks is more prominent during bullish markets, Lin (2018) documents an unproportionate increase in herding activity for small-caps subject to increased uncertainty, and Roger et al. (2018) find analysts to be more optimistic about small price stocks than large price stocks due to inherent cognitive biases in number processing. Additionally, Kremer and Nautz (2013) report herding depends on past stock returns justifying the subsample test based on momentum sorts. Sector-wise, Kim and Pantzalis (2003) show that herding activity is more pronounced for companies that are geographically or industrially diversified.

This study utilises logarithmic forecast errors $ln\left|\frac{\overline{P}_{it}^{A}}{P_{it+1}}-1\right|$ to partially address the impact that outliers could have on the stability of regression coefficients. As additional robustness checks, it also applies quantile regression as in Koenker and Basset (1978) to estimate conditional medians for the full sample and across all subsamples as well as a range of conditional quantiles for the baseline estimation. Further, this study considers non-parametric Spearman rank and Kendall's tau correlation tests for the number of analysts and prediction error scaled upwards by the square root of analyst coverage. If, when adjusted by the square root of the number of analysts, the prediction error shows a positive (negative) non-parametric correlation with coverage, the null hypothesis of independence can be rejected in favour of the alternative hypothesis of herding (contrarianism).

Endogeneity can be a concern, as both the number of analysts following the stock and its forecast error can be influenced by unobserved omitted variables. If a stock is difficult to forecast and analysts value individually accurate forecasts, fewer analysts might be willing to follow such a stock, which would bias $\hat{\beta}$ downwards and leading to false negatives for herding and false positives for contrarianism. Additionally, stocks from various sectors and those adhering to particular investment styles could attract disproportionate attention from analysts due to their specialisation or preferences. For example, O'Brien and Bhushan (1990) report that analysts are more likely to follow industries with a growing number of firms as well as regulated industries. Clement and Tse (2005) show that analysts concentrating on a few industries and having more experience covering similar stocks are more likely to issue bold and accurate forecasts, which might incentivise analysts to specialise narrowly. This finds some reinforcement in this study's sample, with notable heterogeneities observed in coverage across sectors and Fama-French sorts. As such, small-caps, mid-caps, and large-caps are followed each year by 1.35, 3.23, and 13.46 analysts on average. The most followed sector is consumer staples, with 10.36 analysts per stock per year, and the least followed are healthcare and funds, with 4.89 and 4.26, respectively. Growth stocks are more popular with analysts than value stocks, with 8.25 on average covering the former and only 5.34 the latter. These heterogeneities, however, are an excellent foundation for instrumental variable construction. Hence, this study resorts to two-stage least squares regressions with the log of average analyst coverage across similar stocks instrumenting for the log of the observed number of analysts. Three separate regressions are estimated, with coverage in the same sector, in the same sector and year, and in the same sector, year, and Fama-French styles. The validity of instrumental variable regressions is assessed using Durbin-Wu-Hausman endogeneity (Nakamura & Nakamura, 1981) and Anderson-Rubin weak instrument (Anderson & Rubin, 1949) diagnostic tests, as recommended in Young (2019).

Finally, this study considers potential spurious herding concerns. Spurious herding can be distinguished from herding proper in the sense that analyst forecasts might be correlated not due to imitation, but coincidentally due to the application of similar valuation models and techniques (Hwang & Salmon, 2004). Alternatively, a test might mistakenly recognise other analyst behavioral biases, such as optimism and conservatism (Blasco et al., 2018). To address the spurious herding criticism, this study utilises insights from prior research on "herding towards factors" (Hwang & Salmon, 2004), organisational psychology regarding the "magic number" of people that could constitute a group (Argenti, 2020; Collins & Poras, 1996), and studies on number processing biases (Roger et al., 2018). First, if the results are more pronounced for the long sides of Fama-French sorts, it can be interpreted as "herding towards factors" by analysts, augmenting the Hwang and Salmon (2004) perspective from the institutional side. Second, if the nature of herding observed is behavioral and not coincidental, it can be suspected that the herding patterns will be more pronounced after a certain breakpoint in the number of analysts that is sufficient to induce the possibility of imitation and groupthink. Such a "magic number" of group members is commonly assessed as being five or five to seven (Argenti, 2020; Collins & Poras, 1996). This can also separate behavioral herding from competition-driven herding as in Hong and Kacperczyk (2010) and Wang et al. (2020): if herding magnifies with the number of analysts, it is behavioral and caused by group psychology, whereas if herding diminishes with increased coverage, the conflict-of-interest explanation is more plausible. Therefore, this study considers estimations for subsamples based on analyst coverage, undertaking the Chow structural shift test (Chow, 1960) to determine whether such a breakpoint exists and if so, the number of analysts it corresponds to. The presence of herding above the breakpoint and absence thereof below it would reinforce the behavioral and non-spurious nature of detected effects. Finally, the study undertakes tests in subsamples based on the initial stock price. If the degree of herding varies in such estimations, the behavioral motivation of herding can be confirmed and linked with the number processing bias established in Roger et al. (2018).

As for the robustness of herding to other behavioral biases, this study also applies the following iterative procedure, sequentially adjusting the errors for optimism and conservatism, and then applying the regression herding test:

$$ln\frac{\overline{P}_{it}^{A}}{P_{it+1}}=\omega+\kappa\, ln\frac{\overline{P}_{it}^{A}}{P_{it+1}}+u_{it}$$

where ω and κ are the estimators of persistent analyst optimism and conservatism, respectively. This equation is estimated using weighted least squares, scaling each observation by the $(N_{it}^{A})^{-\beta}$, starting with $\beta = -0.5$ as per the null hypothesis of no herding. Next, the residuals u_{it} of the model are used in the conventional regression until convergence, in the full sample and across sectors and years for additional robustness:

$$ln\left|e^{u_{it}}-1\right|=\alpha+\beta\ ln\ N_{it}^{A}+\varepsilon_{it}$$

Data and code for all estimation procedures and robustness tests are available upon request from the corresponding author. In the following section, test results are presented and discussed sequentially.

3. Findings and discussion

This section outlines the data this study uses alongside the application of developed statistical tests. Table 2 reports the descriptive statistics for the main variables, while Table 3 reports coverage across sectors, and Figure 2 visualises the full-sample relationship between the number of analysts and absolute prediction error in a scatterplot. Analyst coverage varies substantially across sample stocks from one to 50 analyst valuations per stock per year, allowing to sufficiently execute the regression-based test. Log absolute forecast error demonstrates behaviour close to normality. However, outliers are present and the concerns regarding their impact on the estimations are al-

Table 2. Descriptive statistics

	Number of analysts	Log forecast error	Log absolute error
Mean	6.9982	0.3564	–1.0413
Median	4	0.1792	–1.1212
Minimum	1	–5.8193	–8.1047
Maximum	50	8.1552	8.1549
Standard deviation	7.6036	0.7988	1.6016
Skewness	1.4811	1.5170	0.1603
Kurtosis	1.6589	8.0523	1.2317
Number of observations	12302	12302	12302

Source: Bloomberg.

Table 3. Sectoral data breakdown

Sector	Number of stocks	Stock-year observations	Analyst valuations
Communications	155	845	6,000
Consumer Discretionary	219	1,487	12,164
Consumer Staples	88	671	6,949
Energy	212	1,229	8,357
Financials	269	1,481	11,958
Funds	23	107	456
Health Care	144	756	3,699
Industrials	302	1,993	14,302
Materials	265	1,526	10,576
Real Estate	117	733	3,988
Technology	246	1,251	5,787
Utilities	39	223	1,617

Source: Bloomberg.

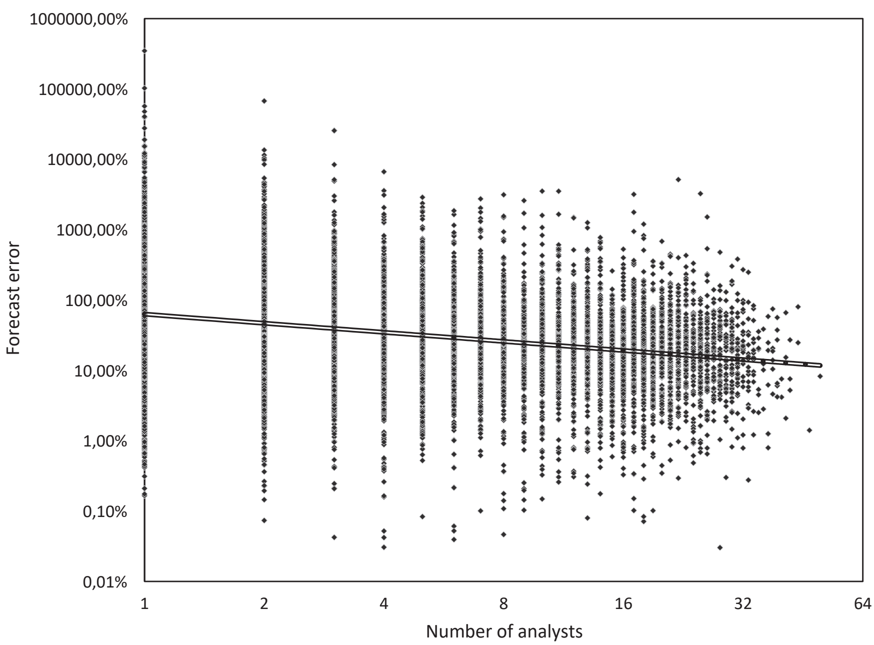

Figure 2. Full sample scatterplot

Source: Bloomberg.

leviated through the non-parametric binomial correlation test and in further robustness checks. Coverage does vary substantially across sectors, allowing for instrumental variable estimations to have sufficiently strong first stages. All the result output tables below report estimated coefficients alongside respective standard errors in parentheses, with ***, **, and * denoting statistical significance at 1%, 5%, and 10%, respectively.

Table 4 reports the binomial correlation test results. There is a significant positive correlation between analyst overvaluations in the full sample, in nine out of twelve sample years, and in most sectors. Herding is not observed for communications, health care, technology, funds, and utilities. This can be explained for the former three sectors, as they are generally technological-

Table 4. Binominal correlation test results

Panel A: Full sample	**Correlation**	**Standard error**	***T*-statistic**	***p*-value**
Across years	0.0449***	(0.0086)	5.2321	0.0000
Across sectors	0.0370***	(0.0086)	4.3066	0.0000
Across years and sectors	0.1188***	(0.0085)	13.9197	0.0000
Panel B: Individual years	**Correlation**	**Standard error**	***T*-statistic**	***p*-value**
2009	0.0562*	(0.0296)	1.8966	0.0581
2010	0.0305	(0.0293)	1.0410	0.2981
2011	0.0294	(0.0289)	1.0171	0.3093
2012	0.1288***	(0.0282)	4.5674	0.0000
2013	0.1778***	(0.0288)	6.1699	0.0000
2014	0.0973***	(0.0298)	3.2676	0.0011
2015	0.0813***	(0.0295)	2.7537	0.0066
2016	0.1010***	(0.0296)	3.4141	0.0007
2017	0.0846***	(0.0302)	2.8025	0.0052
2018	0.0471	(0.0302)	1.5567	0.1198
2019	0.0888***	(0.0306)	2.9009	0.0038
2020	0.0718**	(0.0319)	2.2549	0.0244
Panel C: Individual sectors	**Correlation**	**Standard error**	***T*-statistic**	***p*-value**
Communications	0.0542	(0.0329)	1.6457	0.1002
Consumer Discretionary	0.1301***	(0.0252)	5.1595	0.0000
Consumer Staples	0.1241***	(0.0376)	3.2989	0.0010
Energy	0.1012***	(0.0273)	3.7112	0.0002
Financials	0.1083***	(0.0250)	4.3313	0.0000
Funds	0.1480	(0.0922)	1.6043	0.1114
Health Care	0.0425	(0.0355)	1.2715	0.2039
Industrials	0.0746***	(0.0214)	3.4801	0.0005
Materials	0.1054***	(0.0241)	4.3669	0.0000
Real Estate	0.1733***	(0.0351)	4.9359	0.0000
Technology	0.0334	(0.0256)	1.3056	0.1919
Utilities	0.0627	(0.0642)	0.9771	0.3295

Notes: standard errors reported in parentheses; ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively.

Source: Bloomberg.

ly intensive and more difficult to value, enabling bold contrarian forecasts, and for the latter two, as funds and utilities have stable business models and simple valuation heuristics, yielding public information suitable for a forecast without imitating prior analysts.

The baseline regression test results presented below in Table 5 are generally consistent with these of the binomial correlation test. The null hypothesis of independent valuations is rejected in favour of the alternative hypothesis

Table 5. Regression test results across years and sectors

Panel A: Full sample	**Correlation**	**Standard error**	***T*-statistic**	***p*-value**
Full sample	–0.4291***	(0.0120)	5.9197	0.0000
Panel B: Individual years	**Correlation**	**Standard error**	***T*-statistic**	***p*-value**
2009	–0.4629	(0.0398)	0.9332	0.3509
2010	–0.4346*	(0.0355)	1.8416	0.0658
2011	–0.3273***	(0.0394)	4.3829	0.0000
2012	–0.4929	(0.0387)	0.1824	0.8553
2013	–0.3464***	(0.0364)	4.2245	0.0000
2014	–0.4141**	(0.0431)	1.9926	0.0466
2015	–0.3942**	(0.0454)	2.3305	0.0200
2016	–0.4976	(0.0402)	0.0597	0.9524
2017	–0.4150**	(0.0410)	2.0752	0.0382
2018	–0.4508	(0.0433)	1.1363	0.2561
2019	–0.5728	(0.0473)	–1.5376	0.1245
2020	–0.3796***	(0.0466)	2.5850	0.0099
Panel C: Individual sectors	**Correlation**	**Standard error**	***T*-statistic**	***p*-value**
Communications	–0.4339	(0.0434)	1.5248	0.1277
Consumer Discretionary	–0.3330***	(0.0347)	4.8122	0.0000
Consumer Staples	–0.4276	(0.0490)	1.4790	0.1396
Energy	–0.5252	(0.0398)	–0.6338	0.5263
Financials	–0.2608***	(0.0303)	7.9086	0.0000
Funds	–0.3321	(0.2464)	0.6814	0.4971
Health Care	–0.7693***	(0.0506)	–5.3266	0.0000
Industrials	–0.3855***	(0.0281)	4.0706	0.0000
Materials	–0.5008	(0.0347)	–0.0236	0.9811
Real Estate	–0.2696***	(0.0507)	4.5399	0.0000
Technology	–0.4149**	(0.0382)	2.2306	0.0259
Utilities	–0.7389***	(0.0813)	–2.9380	0.0037

Notes: standard errors reported in parentheses; ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively.

Source: Bloomberg.

of herding for the full sample, for seven out of twelve years, and five out of twelve sectors. Herding is most prominent for financials and real estate, potentially supporting the conflict-of-interest theory (Lee & Lee, 2015) or the groupthink theory.

Next, the prominence of herding is studied across Fama-French portfolio sorts. This is presented in Table 6. Analysts are rational, herding, ad contrarian for low-, medium- and high-beta stocks, respectively. This reiterates the "herding towards beta" concept of Hwang and Salmon (2004) and could also be explained in the context of the "betting against beta" investment strategy

Table 6. Regression test results across Fama-French portfolio sorts

Panel A: Market beta	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Low-beta (bottom 30%)	–0.4552	(0.0368)	1.2200	0.2226
Mid-beta (middle 40%)	–0.4136***	(0.0187)	4.6122	0.0000
High-beta (top 30%)	–0.5789***	(0.0228)	–3.4606	0.0005
Panel B: Size	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Small (bottom 30%)	0.0959***	(0.0932)	6.3930	0.0000
Mid (middle 40%)	–0.1024***	(0.0285)	13.9615	0.0000
Large (top 30%)	–0.1429***	(0.0235)	15.2180	0.0000
Panel C: Value	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Value (top 30%)	–0.4014***	(0.0264)	3.7319	0.0002
Mid (middle 40%)	–0.3670***	(0.0189)	7.0371	0.0000
Growth (bottom 30%)	–0.5183	(0.0213)	–0.8583	0.3908
Panel D: Momentum	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Winner (top 30%)	–0.3069***	(0.0195)	9.8887	0.0000
Mid (middle 40%)	–0.3445***	(0.0171)	9.1052	0.0000
Loser (bottom 30%)	–0.5318	(0.0261)	–1.2146	0.2246
Panel E: Profitability	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Robust (top 30%)	–0.2537***	(0.0201)	12.2532	0.0000
Mid (middle 40%)	–0.2523***	(0.0172)	14.3972	0.0000
Weak (bottom 30%)	–0.4556	(0.0315)	1.4075	0.1594
Panel F: Investment	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Conservative (bottom 30%)	–0.5128	(0.0241)	–0.5326	0.5944
Mid (middle 40%)	–0.3399***	(0.0175)	9.1350	0.0000
Aggressive (top 30%)	–0.4212***	(0.0223)	3.5300	0.0004

Notes: standard errors reported in parentheses; *** denotes statistical significance at 1%.

Source: Bloomberg.

(Frazzini & Pedersen, 2014). Analysts herd for all three sorts based on market capitalisation, although the magnitude of herding is smaller for large-caps. For value, momentum, and profitability, analysts are rational for the short sides of Fama-French factors (growth stocks, past year losers, and weak operating profitability companies) and herd towards their long sides, reinforcing the non-spurious nature of the detected phenomenon. For the investment sorts, the pattern is reversed, which can be explained by lower prominence and institutional reliance on the conservative-minus-aggressive factor documented in Shanaev and Ghimire (2021).

Figure 3 reports Chow structural shift *F*-statistic for candidate breakpoints in terms of the number of analysts. This estimation strategy allows econo-

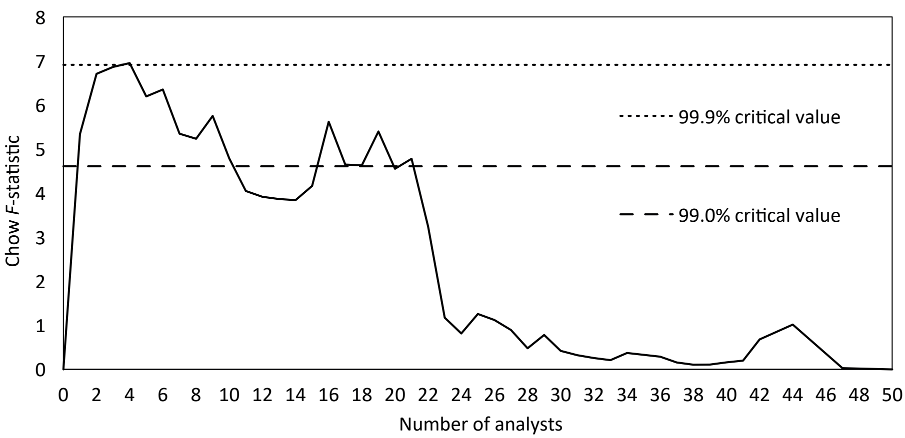

Figure 3. Chow structural break test for number of analysts

Source: Bloomberg.

mists, to distinguish between the conflict-of-interest explanation of herding, which predicts herding will be weaker for more followed stocks due to higher competition (Lee & Lee, 2015), and the behavioral theory relying on group psychology and hypothesising herding will become prominent, starting from a certain number of analysts covering the stock. The strongest structural break is observed for $N_{it}^{A} = 5$, demonstrating the relationship between coverage and log absolute prediction error is different for stocks followed by less than five analysts and by five analysts or higher. Interestingly, five is a commonly cited "magic number" for group formation (Collins & Poras, 1996), providing some early support for the behavioral group psychology explanation. This result dictates the methodological choice of estimating the baseline regression equation for subsamples based on the five analysts breakpoint.

Table 7 reports additional robustness checks for subsamples for varying analyst coverage (a breakpoint of five was chosen based on prior Chow structural shift test estimations) and stock price magnitude. The results support the behavioral nature of analyst herding, since the beta coefficient becomes significantly higher than –0.5 only when the stock is followed by at least five analysts, which is consistent with the group psychology theory and contradicting the conflict-of-interest and competition hypothesis; and for low- and mid-price stocks, supporting the number-processing bias (Roger et al., 2018).

Table 7. Regression test results: Additional robustness checks

Panel A: # of analysts	Coefficient	Standard error	*T*-statistic	*p*-value
Lower than five	–0.5380	(0.0370)	–1.0249	0.3055
Five or higher	–0.3530***	(0.0338)	4.3449	0.0000
Panel B: Stock price	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Low (below £1)	–0.2678***	(0.0334)	6.9414	0.0000
Mid (between £1 and £5)	–0.2773***	(0.0187)	9.9422	0.0000
High (above £5)	–0.4750	(0.0244)	1.0261	0.3049

Notes: standard errors reported in parentheses; *** denotes statistical significance at 1%.

Source: Bloomberg.

To further address heterogeneity and endogeneity bias concerns, Table 8 below reports estimation results in panel regressions with fixed effects and in TSLS regressions, where average coverage across sector, sector and year, and sector, year, and Fama-French factor sorts is instrumenting for the number of analysts, overwhelmingly reinforcing prior findings.

Table 8. Regression test results: Fixed effects and instrumental variable estimations

Panel A: Fixed effects	Coefficient	Standard error	*T*-statistic	*p*-value
Year effects	–0.4303***	(0.0119)	5.8457	0.0000
Sector effects	–0.4231***	(0.0119)	6.4442	0.0000
Sector and year effects	–0.4244***	(0.0119)	6.3687	0.0000
Stock effects	0.1153***	(0.0299)	20.5954	0.0000
Stock and year effects	0.1080***	(0.0302)	20.1474	0.0000
Panel B: IV estimations	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Sector	–0.3023***	(0.0632)	3.1278	0.0018
Sector, year	–0.2773***	(0.0598)	3.7215	0.0002
Sector, year, factor sorts	–0.4528***	(0.0127)	3.7175	0.0002

Notes: standard errors reported in parentheses; *** denotes statistical significance at 1%.

Source: Bloomberg.

The validity and informational value of the instrumental variable estimations discussed above is reinforced with Durbin-Wu-Hausman and Anderson-Rubin tests. It shows that TSLS estimators are significantly different from their OLS counterparts while having very strong first stages. This can be noted in Table 9.

Table 9. Diagnostic tests for instrumental variable estimations

IV estimation	Durbin-Wu-Hausman test		Anderson-Rubin test	
	T-statistic	*p*-value	*T*-statistic	*p*-value
Sector	4.2214**	0.0399	484.61***	0.0000
Sector, year	6.8055***	0.0091	545.36***	0.0000
Sector, year, factor sorts	50.2590***	0.0000	165827.90***	0.0000

Notes: *** and ** denotes statistical significance at 1% and 5%, respectively.

Source: Bloomberg.

Figure 4 reports the herding beta coefficient estimator alongside respective 90% confidence intervals in a quantile regression framework. Analysts are consistently herding when prediction error is low and around the median (up

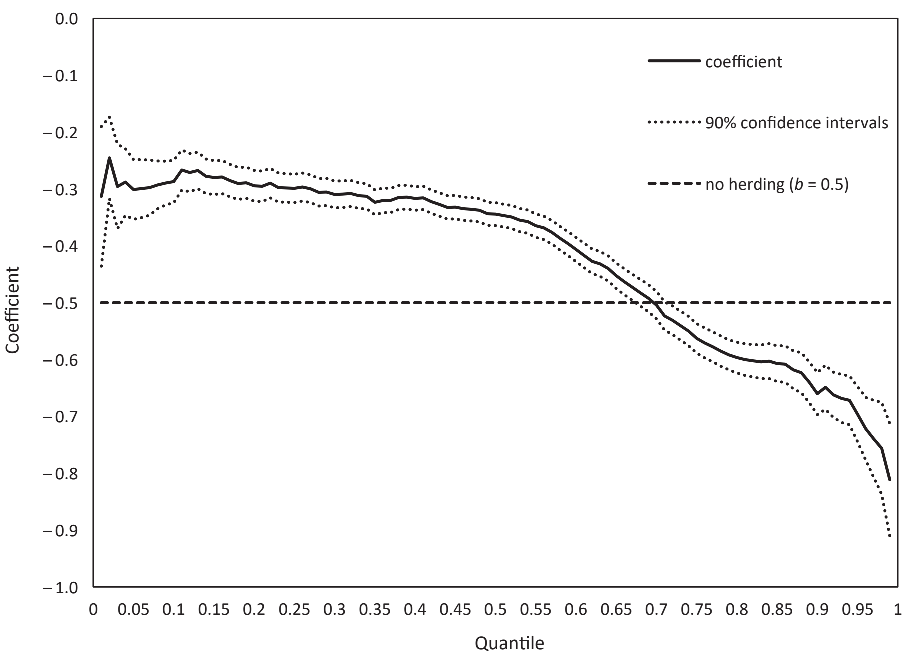

Figure 4. Quantile regression estimation results for the full sample

Source: Bloomberg.

to the 67th percentile), and are, on average, rational from the 68th until the 72nd percentiles and contrarian when forecast error is high (73rd percentile and above). These findings support Hwang and Salmon (2004), who report more prominent herding in investor behaviour during calm market periods, and Welch (2000), who finds a higher degree of analyst herding when market conditions are favourable, while contradicting Caparrelli et al. (2004) and Lin (2018).

To further account for the potential impact of outliers, all subsample-based tests are run in a quantile regression framework for conditional median estimations, with the results presented in Table 10. Consistent with previous findings, herding is confirmed in the full sample in all years except 2019, in all sectors apart from energy and funds (where analysts demonstrate rational behaviour), and utilities (where they are contrarian). The tendency of analysts to herd towards the long sides of Fama-French sorts (apart from the investment factor) and when coverage is sufficient also persists. An additional robustness check revolving around non-parametric Spearman and Kendall's tau correlation coefficients between analyst coverage and adjusted absolute prediction error also largely corroborates the previous results (see Table 11).

Table 10. Quantile regression conditional median estimations

Panel A: Full sample	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Full sample	–0.3442***	(0.0122)	12.7248	0.0000
Panel B: Individual years	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
2009	–0.3320***	(0.0399)	4.2061	0.0000
2010	–0.3153***	(0.0342)	5.4006	0.0000
2011	–0.3105***	(0.0491)	3.8588	0.0001
2012	–0.4254*	(0.0437)	1.7066	0.0882
2013	–0.2428***	(0.0291)	8.8491	0.0000
2014	–0.4058*	(0.0523)	1.7997	0.0722
2015	–0.3170***	(0.0505)	3.6270	0.0003
2016	–0.3610***	(0.0472)	2.9440	0.0033
2017	–0.3043***	(0.0403)	4.8505	0.0000
2018	–0.3521***	(0.0464)	3.1888	0.0015
2019	–0.4279	(0.0512)	1.4095	0.1590
2020	–0.3405***	(0.0529)	3.0177	0.0026
Panel C: Individual sectors	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Communications	–0.3565***	(0.0460)	3.1207	0.0019
Consumer Discretionary	–0.2506***	(0.0316)	7.8938	0.0000
Consumer Staples	–0.3607***	(0.0528)	2.6404	0.0085
Energy	–0.5578	(0.0505)	–1.1440	0.2529
Financials	–0.2239***	(0.0320)	8.6288	0.0000

Funds	–0.4681	(0.2299)	0.1390	0.8897
Health Care	–0.7138***	(0.0670)	–3.1927	0.0015
Industrials	–0.2645***	(0.0264)	8.9114	0.0000
Materials	–0.4894	(0.0423)	0.2514	0.8016
Real Estate	–0.2715***	(0.0538)	4.2444	0.0000
Technology	–0.3465***	(0.0408)	3.7604	0.0002
Utilities	–0.6840*	(0.0946)	–1.9449	0.0531
Panel D: Market beta	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Low-beta (bottom 30%)	–0.3795***	(0.0424)	2.8412	0.0045
Mid-beta (middle 40%)	–0.3362***	(0.0197)	8.2929	0.0000
High-beta (top 30%)	–0.4893	(0.0219)	0.4886	0.6252
Panel E: Size	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Small (bottom 30%)	0.0699***	(0.1225)	4.6512	0.0000
Mid (middle 40%)	–0.0712***	(0.0285)	15.0249	0.0000
Large (top 30%)	–0.0991***	(0.0218)	18.3617	0.0000
Panel F: Value	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Value (top 30%)	–0.3778***	(0.0322)	3.7932	0.0002
Mid (middle 40%)	–0.2976***	(0.0194)	10.4448	0.0000
Growth (bottom 30%)	–0.3973***	(0.0209)	4.9139	0.0000
Panel G: Momentum	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Winner (top 30%)	–0.2434***	(0.0196)	13.1110	0.0000
Mid (middle 40%)	–0.2728***	(0.0178)	12.7876	0.0000
Loser (bottom 30%)	–0.6017***	(0.0321)	–3.1651	0.0016
Panel H: Profitability	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Robust (top 30%)	–0.1971***	(0.0201)	15.0565	0.0000
Mid (middle 40%)	–0.1958***	(0.0166)	18.3502	0.0000
Weak (bottom 30%)	–0.4750	(0.0383)	0.6532	0.5137
Panel I: Investment	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Conservative (bottom 30%)	–0.4627	(0.0271)	1.3782	0.1682
Mid (middle 40%)	–0.2715***	(0.0172)	13.2478	0.0000
Aggressive (top 30%)	–0.3362***	(0.0242)	6.7675	0.0000
Panel J: Number of analysts	**Coefficient**	**Standard error**	***T*-statistic**	***p*-value**
Lower than five	–0.4650	(0.0413)	0.8466	0.3972
Five or higher	–0.2445***	(0.0320)	7.9931	0.0000

Notes: standard errors reported in parentheses; *** and * denote statistical significance at 1% and 10%, respectively.

Source: Bloomberg.

Table 11. Non-parametric correlation tests

Panel A: Full sample	**Spearman**	***p*-value**	**Kendall's tau**	***p*-value**
Full sample	0.0743***	0.0000	0.0544***	0.0000
Panel B: Individual years	**Spearman**	***p*-value**	**Kendall's tau**	***p*-value**
2009	0.0713**	0.0198	0.0538**	0.0128
2010	0.1089***	0.0003	0.0790***	0.0002
2011	0.1407***	0.0000	0.0986***	0.0000
2012	0.0132	0.6561	0.0120	0.5657
2013	0.1927***	0.0000	0.1432***	0.0000
2014	0.0752**	0.0183	0.0550**	0.0138
2015	0.0755**	0.0164	0.0588***	0.0078
2016	0.0287	0.3599	0.0251	0.2552
2017	0.1027***	0.0011	0.0769***	0.0005
2018	0.0597*	0.0587	0.0439**	0.0476
2019	–0.0245	0.4463	–0.0149	0.5075
2020	0.0992***	0.0031	0.0695***	0.0030
Panel C: Individual sectors	**Spearman**	***p*-value**	**Kendall's tau**	***p*-value**
Communications	0.0666*	0.0529	0.0505**	0.0376
Consumer Discretionary	0.1660***	0.0000	0.1175***	0.0000
Consumer Staples	0.0727*	0.0597	0.0510*	0.0539
Energy	–0.0096	0.7371	–0.0078	0.6963
Financials	0.2367***	0.0000	0.1683***	0.0000
Funds	0.0577	0.5546	0.0400	0.5769
Health Care	–0.1669***	0.0000	–0.1198***	0.0000
Industrials	0.1313***	0.0000	0.0931***	0.0000
Materials	0.0074	0.7731	0.0075	0.6788
Real Estate	0.1744***	0.0000	0.1241***	0.0000
Technology	0.0662**	0.0193	0.0510**	0.0126
Utilities	–0.1925***	0.0039	–0.1424***	0.0027
Panel D: Market beta	**Spearman**	***p*-value**	**Kendall's tau**	***p*-value**
Low-beta (bottom 30%)	0.0306	0.1035	0.0230	0.1006
Mid-beta (middle 40%)	0.0823***	0.0000	0.0599***	0.0000
High-beta (top 30%)	–0.0173	0.2601	–0.0081	0.4386
Panel E: Size	**Spearman**	***p*-value**	**Kendall's tau**	***p*-value**
Small (bottom 30%)	0.1198***	0.0000	0.0953***	0.0000
Mid (middle 40%)	0.1985***	0.0000	0.1449***	0.0000
Large (top 30%)	0.2291***	0.0000	0.1587***	0.0000
Panel F: Value	**Spearman**	***p*-value**	**Kendall's tau**	***p*-value**
Value (top 30%)	0.0664***	0.0003	0.0484***	0.0002
Mid (middle 40%)	0.1223***	0.0000	0.0880***	0.0000
Growth (bottom 30%)	0.0255	0.1328	0.0208*	0.0768
Panel G: Momentum	**Spearman**	***p*-value**	**Kendall's tau**	***p*-value**
Winner (top 30%)	0.1980***	0.0000	0.1411***	0.0000

Mid (middle 40%)	0.1495***	0.0000	0.1066***	0.0000
Loser (bottom 30%)	–0.0272	0.1285	–0.0189	0.1387
Panel H: Profitability	**Spearman**	***p*-value**	**Kendall's tau**	***p*-value**
Robust (top 30%)	0.2308***	0.0000	0.1599***	0.0000
Mid (middle 40%)	0.2288***	0.0000	0.1624***	0.0000
Weak (bottom 30%)	0.0232	0.2349	0.0172	0.2258
Panel I: Investment	**Spearman**	***p*-value**	**Kendall's tau**	***p*-value**
Conservative (bottom 30%)	–0.0064	0.7217	–0.0025	0.8476
Mid (middle 40%)	0.1603***	0.0000	0.1135***	0.0000
Aggressive (top 30%)	0.0735***	0.0000	0.0534***	0.0000
Panel J: Number of analysts	**Spearman**	***p*-value**	**Kendall's tau**	***p*-value**
Lower than five	–0.0057	0.6340	–0.0043	0.6364
Five or higher	0.0789***	0.0000	0.0548***	0.0000

Notes: ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively.

Source: Bloomberg.

Table 12 presents the results of an iterative estimation adjusted for analyst optimism and conservatism. The security analysts have been excessively optimistic in all years apart from 2012 and 2019, when they were rational, and 2013, when they were overly pessimistic; and in all sectors apart from funds that are naturally easiest to value, and significantly conservative in all subsamples. Analysts are shown to consistently overestimate target forecast prices by 14% on average, which is consistent with the figures reported previously in the literature (Bradshaw, Brown et al., 2013; Brav & Lehavy, 2003). The results for herding when other prominent behavioral biases are accounted for become more pronounced, with the beta coefficient significantly higher than -0.5 in all sample years and in all sectors but funds and utilities. This robustness test reinforces the non-spurious nature of the herding detected by the regression test and illuminates the relationship between analyst herding and other biases that manifest in their valuations.

Conclusions

This study has developed two novel, flexible, and powerful tests based on binominal correlations and prediction error volatility scaling for herding in analyst target forecast prices. The purpose is to reinforce the existence and prominence of herding patterns among analysts observed over the 2008–2020 sample period in the United Kingdom. It contributes substantially to the de-

Table 12. Regression test results adjusted for analyst optimism and conservatism

Panel A: Full sample	Optimism (ω)	Standard error	Conservatism ($\varkappa$)	Standard error	Herding (β)	Standard error
Baseline	0.1404***	(0.0064)	0.5513***	(0.0188)	–0.2159***	(0.0116)
Panel B: Individual years	**Optimism (ω)**	**Standard error**	**Conservatism ($\varkappa$)**	**Standard error**	**Herding (β)**	**Standard error**
2010	0.0365**	(0.0180)	0.3894***	(0.0776)	–0.3060***	(0.0358)
2011	0.4142***	(0.0198)	0.5473***	(0.0698)	–0.1666***	(0.0314)
2012	0.0227	(0.0321)	0.5219***	(0.0716)	–0.2161***	(0.0385)
2013	–0.0959***	(0.0159)	0.7262***	(0.0439)	–0.2449***	(0.0340)
2014	0.2994***	(0.0187)	0.6927***	(0.0553)	–0.1998***	(0.0372)
2015	0.1107***	(0.0174)	0.8463***	(0.0461)	–0.2348***	(0.0417)
2016	0.0374*	(0.0277)	0.4691***	(0.0648)	–0.2410***	(0.0385)
2017	0.0410**	(0.0183)	0.4900***	(0.0697)	–0.2417***	(0.0405)
2018	0.3476***	(0.0153)	0.5823***	(0.0545)	–0.2867***	(0.0392)
2019	–0.0137	(0.0215)	0.7738***	(0.0571)	–0.2498***	(0.0429)
2020	0.1465***	(0.0160)	0.5647***	(0.0469)	–0.1806***	(0.0451)
Panel C: Individual sectors	**Optimism (ω)**	**Standard error**	**Conservatism ($\varkappa$)**	**Standard error**	**Herding (β)**	**Standard error**
Communications	0.1265***	(0.0201)	0.5131***	(0.0560)	–0.2219***	(0.0429)
Consumer Discretionary	0.1350***	(0.0154)	0.3975***	(0.0523)	–0.1760***	(0.0333)
Consumer Staples	0.0866***	(0.0150)	0.3822***	(0.0501)	–0.2553***	(0.0490)
Energy	0.3208***	(0.0255)	0.5570***	(0.0397)	–0.2675***	(0.0358)
Financials	0.1046***	(0.0126)	0.4627***	(0.0678)	–0.0924***	(0.0310)
Funds	0.0298	(0.0434)	0.5798***	(0.2196)	–0.4557	(0.3104)
Health Care	0.1433***	(0.0230)	0.6075***	(0.0552)	–0.3593***	(0.0454)
Industrials	0.1088***	(0.0117)	0.4900***	(0.0450)	–0.2271***	(0.0278)
Materials	0.1737***	(0.0259)	0.6577***	(0.0462)	–0.2089***	(0.0311)
Real Estate	0.1491***	(0.0212)	0.3683***	(0.0731)	–0.1361***	(0.0509)
Technology	0.1142***	(0.0169)	0.3900***	(0.0463)	–0.2030***	(0.0409)
Utilities	0.0630**	(0.0290)	0.4946***	(0.1211)	–0.5918	(0.0819)

Notes: standard errors reported in parentheses; ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively.

Source: Bloomberg.

bates on the presence, origins, and conditions of institutional herding prominent in the existing literature.

As such, the findings of the study confirm that herding is non-spurious and behavioral in nature, as analyst herding is more prominent towards the long sides of the Fama-French factor portfolio sorts (small and value stocks, past year winners, and companies with high operating profitability), consistent with the betting-against-beta strategy, exacerbates when a stock is followed by at least five analysts, corresponding to the insights from group psychology, and for low- and mid-price stocks, reinforcing the number processing bias hypothesis. The results are largely inconsistent with the conflict-of-interest explanation and the competition hypothesis, as herding does not diminish with increased analyst coverage. Subsample and conditional quantile estimations show that herding is more prominent when uncertainty and market volatility is low. The flexibility of the derived econometric tests allows them to be applied to aggregated, readily available data, and to simultaneously test for herding, optimism, and conservatism in an iterative regression framework, while also not requiring specialised, high-frequency, or analyst-level datasets.

For practitioners and policymakers, the study has confirmed that analyst forecasts, while potentially yielding informational value, are affected by mostly behavioral rather than institutionally driven biases, thus herding in security analysts might not be as easy to address with policy interventions, incentive design, or governance practices as previously thought. Furthermore, as herding is more prominent during calm market periods, its contribution towards market fragility and systemic risk can be found lower than presumed. Individual investors could use the findings of this study to assess the reliability of the analyst consensus for various stocks subject to different market conditions.

The validity and consistency of the obtained results is evidenced across both tests as well as being subject to a battery of robustness checks accounting for heterogeneity and endogeneity biases and also the effect of outliers. Findings persist in 1) sectoral and yearly subsamples; 2) in panel regressions with cross-sectional (sectoral and firm-level) and time fixed effects; 3) two--stage least squares estimations with average coverage across similar stocks instrumenting for the number of analysts observed; 4) conditional median models in the quantile regression framework; 5) in non-parametric correlation tests; and 6) when adjusted for other behavioral biases, such as optimism (pessimism) and conservatism (recency bias). Future research could apply the procedures derived in this study to other security analyst forecasts, such as earnings estimates, and also test for the robustness of this study's results on other prominent international markets, such as the United States, Japan, or the European Union.

References

Abarbanell, J., & Lehavy, R. (2003). Biased forecasts or biased earnings? The role of reported earnings in explaining apparent bias and over/underreaction in analysts' earnings forecasts. *Journal of Accounting and Economics*, *36*(1–3), 105–146.

Aharon, D. Y. (2021). Uncertainty, fear and herding behavior: Evidence from size-ranked portfolios. *Journal of Behavioral Finance*, 22(3), 320–337.

Anderson, T., & Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics*, *20*(1), 46–63.

Argenti, P. (2020, March 13). Communicating through the coronavirus crisis. *Harvard Business Review*.

Bernhardt, D., Campello, M., & Kutsoati, E. (2006). Who herds? *Journal of Financial Economics*, *80*(3), 657–675.

Blake, D., Sarno, L., & Zinna, G. (2017). The market for lemmings: The herding behavior of pension funds. *Journal of Financial Markets*, *36*(1), 17–39.

Blasco, N., Corredor, P., & Ferrer, E. (2018). Analysts herding: When does sentiment matter? *Applied Economics*, *50*(51), 5495–5509.

Bradshaw, M., Brown, L., & Huang, K. (2013). Do sell-side analysts exhibit differential target price forecasting ability? *Review of Accounting Studies*, *18*(4), 930–955.

Bradshaw, M., Huang, A., & Tan, H. (2014). *Analyst target price optimism around the world.* Working Paper, Boston College.

Brav, A., & Lehavy, R. (2003). An empirical analysis of analysts' target prices: Short-term informativeness and long-term dynamics. *Journal of Finance*, *58*(5), 1933–1967.

Caparrelli, F., D'Arcangelis, A., & Cassuto, A. (2004). Herding in the Italian stock market: A case of behavioral finance. *Journal of Behavioral Finance*, *5*(4), 222–230.

Chang, E., Cheng, J., & Khorana, A. (2000). An examination of herd behavior in equity markets: An international perspective. *Journal of Banking & Finance*, *24*(10), 1651–1679.

Chen, Q., & Jiang, W. (2006). Analysts' weighting of private and public information. *Review of Financial Studies*, *19*(1), 319–355.

Cheng, L., Su, Y., Yan, Z., & Zhao, Y. (2019). Corporate governance and target price accuracy. *International Review of Financial Analysis*, *64*(1), 93–101.

Choi, N., & Sias, R. (2009). Institutional industry herding. *Journal of Financial Economics*, *94*(3), 469–491.

Choi, N., & Skiba, H. (2015). Institutional herding in international markets. *Journal of Banking & Finance*, *55*(1), 246–259.

Chow, G. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society*, *28*(3), 591–605.

Christie, W., & Huang, R. (1995). Following the pied piper: Do individual returns herd around the market? *Financial Analysts Journal*, *51*(4), 31–37.

Clement, M. B., & Tse, S. Y. (2003). Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters?. *The Accounting Review*, *78*(1), 227–249.

Clement, M., & Tse, S. (2005). Financial analyst characteristics and herding behavior in forecasting. *Journal of Finance*, *60*(1), 307–341.

Collins, J. & Porras, J. (1996). Building your company's vision. *Harvard Business Review*, *74*(5), 65–77.

Cote, J., & Goodstein, J. (1999). A breed apart? Security analysts and herding behavior. *Journal of Business Ethics*, *18*(3), 305–314.

De Bondt, W., & Forbes, W. (1999). Herding in analyst earnings forecasts: Evidence from the United Kingdom. *European Financial Management*, *5*(2), 143–163.

Doukas, J., Kim, C., & Pantzalis, C. (2005). The two faces of analyst coverage. *Financial Management*, *34*(2), 99–125.

Drehmann, M., Oechssler, J., & Roider, A. (2005). Herding and contrarian behavior in financial markets: An internet experiment. *American Economic Review*, *95*(5), 1403–1426.

Fama, E., & French, K. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22.

Fama, E., & French, K. (2018). Choosing factors. *Journal of Financial Economics*, *128*(2), 234–252.

Feldman, R., Livnat, J., & Zhang, Y. (2012). Analysts' earnings forecast, recommendation, and target price revisions. *Journal of Portfolio Management*, *38*(3), 120–132.

Frazzini, A., & Pedersen, L. (2014). Betting against beta. *Journal of Financial Economics*, *111*(1), 1–25.

Friesen, G., & Weller, P. A. (2006). Quantifying cognitive biases in analyst earnings forecasts. *Journal of Financial Markets*, *9*(4), 333–365.

Frijns, B., & Huynh, T. (2018). Herding in analysts' recommendations: The role of media. *Journal of Banking & Finance*, *91*(1), 1–18.

Galariotis, E., Rong, W., & Spyrou, S. (2015). Herding on fundamental information: A comparative study. *Journal of Banking & Finance*, *50*(1), 589–598.

Graham, J. (1999). Herding among investment newsletters: Theory and evidence. *Journal of Finance*, *54*(1), 237–268.

Guo, J., Holmes, P., & Altanlar, A. (2020). Is herding spurious or intentional? Evidence from analyst recommendation revisions and sentiment. *International Review of Financial Analysis*, *71*(1), 101539.

Han, C., Kang, J., & Kim, S. (2021). Betting against analyst target price. *Journal of Financial Markets*, *59*, 100677.

Hong, H., & Kacperczyk, M. (2010). Competition and bias. *The Quarterly Journal of Economics*, *125*(4), 1683–1725.

Huang, J., Mian, G., & Sankaraguruswamy, S. (2009). The value of combining the information content of analyst recommendations and target prices. *Journal of Financial Markets*, *12*(4), 754–777.

Hwang, S., & Salmon, M. (2004). Market stress and herding. *Journal of Empirical Finance*, *11*(4), 585–616.

Imam, S., Chan, J., & Shah, S. (2013). Equity valuation models and target price accuracy in Europe: Evidence from equity reports. *International Review of Financial Analysis*, *28*(1), 9–19.

Ishigami, S., & Takeda, F. (2018). Market reactions to stock rating and target price changes in analyst reports: Evidence from Japan. *Journal of International Financial Markets, Institutions and Money*, *52*(1), 134–151.

James, C., & Karceski, J. (2006). Strength of analyst coverage following IPOs. *Journal of Financial Economics*, *82*(1), 1–34.

JPMorgan (1997). *CreditMetrics – technical document*. JPMorgan.

Kałdoński, M., & Jewartowski, T. (2017). Agency costs of overvalued equity and earnings management in companies listed on WSE. *Economics and Business Review, 3*(1), 7–37.

Kim, C., & Pantzalis, C. (2003). Global/industrial diversification and analyst herding. *Financial Analysts Journal*, *59*(2), 69–79.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica: Journal of the Econometric Society*, *46*(1), 33–50.

Kremer, S., & Nautz, D. (2013). Causes and consequences of short-term institutional herding. *Journal of Banking & Finance*, *37*(5), 1676–1686.

Lee, J., & Lee, J. (2015). Analyst herding behavior and analyst affiliation: Evidence from business groups. *Journal of Behavioral Finance*, *16*(4), 373–386.

Lin, M. (2018). The impact of aggregate uncertainty on herding in analysts' stock recommendations. *International Review of Financial Analysis*, *57*(1), 90–105.

Lin, S., Tan, H., & Zhang, Z. (2016). When analysts talk, do institutional investors listen? Evidence from target price changes. *Financial Review*, *51*(2), 191–223.

Loang, O. K., & Ahmad, Z. (2021). Does volatility mediate the impact of analyst recommendations on herding in Malaysian stock market?. *Economics and Business Review*, *7*(4), 54–71.

Nakamura, A., & Nakamura, M. (1981). On the relationships among several specification error tests presented by Durbin, Wu, and Hausman. *Econometrica: Journal of the Econometric Society*, *49*(6), 1583–1588.

Nofsinger, J., & Sias, R. (1999). Herding and feedback trading by institutional and individual investors. *Journal of Finance*, *54*(6), 2263–2295.

O'Brien, P. (1988). Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics*, *10*(1), 53–83.

O'Brien, P., & Bhushan, R. (1990). Analyst following and institutional ownership. *Journal of Accounting Research*, *28*(1), 55–76.

Olsen, R. (1996). Implications of herding behavior for earnings estimation, risk assessment, and stock returns. *Financial Analysts Journal*, *52*(4), 37–41.

Roger, T., Roger, P., & Schatt, A. (2018). Behavioral bias in number processing: Evidence from analysts' expectations. *Journal of Economic Behavior & Organization*, *149*(1), 315–331.

Roider, A., & Voskort, A. (2016). Reputational herding in financial markets: A laboratory experiment. *Journal of Behavioral Finance*, *17*(3), 244–266.

Scharfstein, D., & Stein, J. (1990). Herd behavior and investment. *American Economic Review*, *80*(3), 465–479.

Shanaev, S., & Ghimire, B. (2021). Efficient scholars: Academic attention and the disappearance of anomalies. *The European Journal of Finance*, *27*(3), 278–304.

Sias, R. (2004). Institutional herding. *Review of Financial Studies*, *17*(1), 165–206.

Trueman, B. (1994). Analyst forecasts and herding behavior. *Review of Financial Studies*, *7*(1), 97–124.

Vidal-Tomas, D., Ibanez, A., & Farinos, J. (2019). Herding in the cryptocurrency market: CSSD and CSAD approaches. *Finance Research Letters*, *30*(1), 181–186.

Wang, Z., Sun, L., & Wei, K. (2020). Does competition induce analyst effort? Evidence from a natural experiment of broker mergers. *Journal of Banking & Finance*, *119*(1), 105914.

Welch, I. (2000). Herding among security analysts. *Journal of Financial Economics*, *58*(3), 369–396.

Wheeler, F. P., Neale, B., Kowalski, T., & Letza, S. R. (2002). The efficiency of the Warsaw Stock Exchange: The first few years 1991–1996. *Economics and Business Review*, *2*(2), 37–58.

Young, A. (2019). *Consistency without inference: Instrumental variables in practical application*. London School of Economics.