Mikko Malkamäki

Herding behavior in Indian and Chinese equity markets during the COVID-19 pandemic

School of Accounting and Finance
Master’s thesis in Finance
Master’s degree Programme in Finance

Vaasa 2026

UNIVERSITY OF VAASA
School of Accounting and Finance
Author: Mikko Malkamäki
Title of the Thesis: Herding behavior in Indian and Chinese equity markets during the COVID-19 pandemic
Degree: Master of Science in Economics and Business Administration
Program: Master’s Degree Programme in Finance
Supervisor: Veda Fatmy
Year: 2026 **Number of pages:** 79

ABSTRACT:
This thesis examines market-wide herding behavior in the Indian and Chinese equity markets during the COVID-19 period. The study examines whether herding was present during the sample period and whether it changed under different market conditions. Given the partly inconclusive empirical evidence on herding in India and China, and on herding during extreme market events in general, the COVID-19 period provides a relevant recent setting for studying herding in these markets during market uncertainty. Furthermore, as India and China have become major global economies and important emerging markets, understanding their stock market behavior has become increasingly important for both domestic and international investors.

The applied methodology in this thesis is the cross-sectional absolute deviation (CSAD) method, which examines market-wide herding by analyzing the relationship between stock return dispersion and market returns. A modified CSAD specification is used in the regression analysis, which accounts for first-order autoregressive serial correlation. The empirical analysis is based on daily stock return data from 2.7.2018 to 30.6.2022. Herding is examined using three regression tests. The first test assesses whether herding is present over the full sample period. The second examines whether herding varies across three subperiods: pre-COVID, outbreak, and post-outbreak. The third investigates whether herding differs between up- and down-market days.

For India, the results do not provide statistically significant evidence of market-wide herding during the full sample period, across the three subperiods, or during up- and down-market days. For China, the results indicate statistically significant herding in the full sample period, during the outbreak period, and during down-market days. These findings suggest that herding in China may have become more pronounced during periods of high uncertainty and falling market conditions. However, the evidence for China is sensitive to model specification, as statistically significant herding is not found when the AR(1) term is excluded from the regression.

The findings suggest that the Indian equity market behaved more consistently with rational asset pricing predictions. For China, the results support the view that herding may become more pronounced under extreme market conditions and may vary with market direction. Future research could extend the analysis to sector-level herding, different investor groups, and alternative herding methodologies. Furthermore, as current herding methodologies have limitations, research could focus on developing improved methods for detecting herding.

Keywords: Behavioral finance, COVID-19, cross-sectional absolute deviation, emerging markets, herding behavior, market-wide herding

VAASAN YLIOPISTO
Laskentatoimen ja rahoituksen yksikkö
Tekijä: Mikko Malkamäki
Tutkielman nimi: Herding behavior in Indian and Chinese equity markets during the COVID-19 pandemic
Tutkinto: Kauppatieteiden maisteri
Oppiaine: Rahoitus
Ohjaaja: Veda Fatmy
Vuosi: 2026 **Sivumäärä:** 79

Tiivistelmä:
Tämä tutkielma tarkastelee markkinanlaajuista laumakäyttäytymistä Intian ja Kiinan osakemarkkinoilla COVID-19-pandemian aikana. Tutkimuksessa selvitetään, esiintyikö laumakäyttäytymistä tarkastelujaksolla ja vaihteliko se eri markkinaolosuhteissa. Koska aikaisempi empiirinen näyttö laumakäyttäytymisestä Intiassa ja Kiinassa sekä äärimmäisten markkinaolosuhteiden aikana on osittain epäselvää, COVID-19-pandemia tarjoaa ajankohtaisen ja merkityksellisen tutkimusasetelman laumakäyttäytymisen tutkimiseen markkinaepävarmuuden aikana. Intiasta ja Kiinasta on muodostunut merkittäviä talouksia ja tärkeitä kehittyviä markkinoita, minkä takia niiden osakemarkkinoiden käyttäytymisen ymmärtäminen on tullut yhä tärkeämmäksi sekä kotimaisille että kansainvälisille sijoittajille.

Tutkielmassa käytetään cross-sectional absolute deviation (CSAD)-menetelmää, jolla arvioidaan markkinoiden laajuista laumakäyttäytymistä tarkastelemalla osakkeiden tuottojen hajonnan ja markkinatuoton välistä suhdetta. Regressioanalyysissä käytetään muokattua CSAD-mallia, joka huomioi ensimmäisen asteen autokorrelaation. Empiirinen analyysi perustuu päivittäisiin osaketuottoihin ajalta 2.7.2018–30.6.2022. Laumakäyttäytymistä tarkastellaan kolmen regressiotestin avulla. Ensimmäisessä testissä arvioidaan, esiintyykö laumakäyttäytymistä koko otosjaksolla. Toisessa testissä tutkitaan, vaihteleeko laumakäyttäytyminen kolmen osajakson välillä: ennen COVID-pandemiaa, puhkeamisjakson (outbreak period) aikana ja puhkeamisjakson jälkeen. Kolmannessa testissä tutkitaan, eroaako laumakäyttäytyminen nousu- ja laskupäivinä.

Intian osalta tulokset eivät anna tilastollisesti merkitsevää näyttöä markkinanlaajuisesta laumakäyttäytymisestä koko otoksen ajanjaksona, kolmen osajakson aikana tai nousu- ja laskupäivinä. Kiinan osalta tulokset osoittavat tilastollisesti merkitsevää laumakäyttäytymistä koko otosjaksolla, puhkeamisjaksolla sekä laskumarkkinapäivinä. Nämä havainnot viittaavat siihen, että laumakäyttytyminen Kiinassa on saattanut voimistua suuren epävarmuuden ja laskevien markkinaolosuhteiden aikana. Kiinaa koskeva näyttö on kuitenkin herkkä mallispesifikaatiolle, koska tilastollisesti merkitsevää laumakäyttäytymistä ei ilmene, kun AR(1)-termi jätetään regressiosta pois.

Tulokset viittaavat siihen, että Intian osakemarkkinat käyttäytyivät johdonmukaisemmin rationaalisten osakkeiden hinnoittelumallien mukaisesti. Kiinan osalta tulokset tukevat näkemystä, jonka mukaan laumakäyttäytyminen voi voimistua äärimmäisissä markkinaolosuhteissa ja vaihdella markkinan suunnan mukaan. Tuleva tutkimus voisi laajentaa analyysiä sektoritason laumakäyttäytymiseen, eri sijoittajaryhmiin sekä vaihtoehtoisiin tutkimusmenetelmiin. Koska nykyisiin tutkimusmenetelmiin liittyy rajoitteita, tuleva tutkimus voisi keskittyä menetelmien kehittämiseen laumakäyttäytymisen havaitsemiseksi.

Avainsanat: Behavioristinen rahoitus, COVID-19, cross-sectional absolute deviation, kehittyvät markkinat, laumakäyttäytyminen, markkinanlaajuinen laumakäyttäytyminen

Table of contents

Figures

Tables

Abbreviations

APT	Arbitrage Pricing Theory
CAPM	Capital Asset Pricing Model
CDT	Classical Decision Theory
CML	Capital Market Line
CSAD	Cross-Sectional Absolute Deviation
CSSD	Cross-Sectional Standard Deviation
ECSAD	Expected Cross-Sectional Absolute Deviation
EMH	Efficient Market Hypothesis
EUT	Expected Utility Theory
MPT	Modern Portfolio Theory
SML	Security Market Line

1 Introduction

Financial markets have traditionally been explained through theories that assume rational investors and efficient markets. However, empirical research suggests that investor behavior often deviates from these assumptions. Behavioral finance has emerged to explain such deviations by emphasizing that actual investor behavior may be influenced by psychological and emotional factors that, in turn, may affect market outcomes (Baker & Ricciardi, 2014). One widely studied phenomenon that may explain these deviations is herding behavior. In the finance literature, herding generally refers to a situation in which investors imitate others' actions or make their decisions based on the observed behavior of others (Spyrou, 2013).

Understanding herding in financial markets is relevant as it may increase volatility, reduce market stability, and cause asset prices to deviate from their fundamental values (Bikhchandani & Sharma, 2001). If herding persists for longer periods, it can even lead to asset price bubbles and to subsequent crashes (Baker & Ricciardi, 2014). Therefore, it is perhaps not surprising that herding is often viewed by the public as a major cause of market instability and extreme market events (Spyrou, 2013).

As emerging markets play an increasingly important role in the global economy, understanding their market behavior has become more relevant. In the literature, herding behavior has often been linked to market inefficiencies, such as weak market regulation, frequent central bank and government intervention, a larger presence of less-informed investors, and lower information disclosure standards for listed firms (Lao & Singh, 2011). These features are commonly associated with developing markets, and some studies suggest that investors in such markets may be more prone to herding (Lao & Singh, 2011; Chang et al., 2000). In emerging market research, India and China are particularly relevant as they have structural characteristics, such as large geographic size, extensive labor resources, government policies supporting economic growth, and a growing middle class, that support sustained long-term economic growth (Lao & Singh, 2011). Furthermore, rapid economic growth has increased the purchasing power and global importance of India and China, making their stock markets increasingly attractive to international investors. Given their growing economic importance, it is essential for investors to better understand the dynamics of these stock markets, including whether investor behavior exhibits herding.

The relevance of studying herding became even more pronounced during the COVID-19 pandemic, which created a period of uncertainty and high volatility in the global financial markets. In herding research, periods of market stress are considered especially important because investors are thought to be more likely to ignore their own information and follow market consensus under such conditions (Christie & Huang, 1995). However, previous findings on herding during crisis periods are not fully consistent, as the literature does not provide consistent evidence on whether extreme market environments strengthen herding behavior (Christie & Huang, 1995; Chiang & Zheng, 2010). As evidence on herding during crisis periods remains mixed, the COVID-19 period provides a useful recent setting for examining whether herding occurred in the Indian and Chinese equity markets during increased uncertainty and market stress.

1.1 Purpose of the study

This thesis examines market-wide herding behavior during the COVID-19 pandemic in the Indian and Chinese equity markets. More specifically, the study investigates whether herding occurs throughout the full sample period, 2.7.2018-30.6.2022, whether it varies across three subperiods, and whether it is asymmetric between up- and down-market days. By focusing on two major emerging markets during an uncertain period of market stress, this thesis aims to provide further evidence of how herding may occur under extreme market conditions.

The motivation for studying herding across different markets and circumstances is based on earlier studies. Previous evidence suggests that herding may vary across markets, between different market conditions, and between rising and falling markets (Chang et al., 2000; Chiang & Zheng, 2010). Therefore, this thesis examines herding from several perspectives. First, examining herding over the full sample period provides a general assessment of whether market-wide herding was present around the COVID-19 pandemic. Second, a subperiod analysis comprising pre-COVID, outbreak, and post-outbreak periods enables the study to assess whether herding differs over time as market conditions change. Third, an analysis of up- and down-market days allows the study to examine whether herding differs across rising and falling markets. Overall, these perspectives provide a broad view of potential herding behavior in the Indian and Chinese equity markets.

1.2 Contribution to existing literature

Although herding has been relatively widely studied, empirical evidence remains mixed and sometimes inconclusive, depending on the market, investor type, and methodology used (Spyrou, 2013). In addition, research on herding in emerging stock markets remains more limited than in major and developed equity markets, which highlights the need for further evidence from emerging markets (Spyrou, 2013; Lao & Singh, 2011). Previous evidence suggests that herding has been present in the Indian stock market (Lao & Singh, 2011; Bhaduri & Mahapatra, 2013), while evidence from the Chinese stock market is more mixed, with some studies reporting herding (Lao & Singh, 2011; Tan et al., 2008; Chiang et al., 2010) and others finding no such evidence (Demirer & Kutan, 2006). In addition, some previous studies suggest that herding in India is more likely to arise under extreme market conditions (Lao & Singh, 2011; Bhaduri & Mahapatra, 2013), and some others suggest the same for China (Tan et al., 2008), making the COVID-19 period a relevant more recent setting for further examination of herding under extreme market conditions in these markets. Furthermore, comparative evidence on the Indian and Chinese equity markets during the COVID-19 period remains limited.

This thesis, therefore, contributes to the literature by providing additional evidence on market-wide herding in the Indian and Chinese equity markets during the COVID-19 period. The study contributes to the partly mixed empirical literature by providing a comparative analysis of whether market-wide herding was present in two major emerging equity markets during a recent period of heightened uncertainty and market stress. The findings extend beyond a simple full-period examination by considering variation across subperiods, different market conditions, and up- and down-market days. Therefore, the results provide more comprehensive evidence of investor behavior during COVID-19 than a simple full-period analysis and may help to better understand how herding can occur under extreme market conditions in India and China.

1.3 Structure of the study

This thesis consists of seven chapters. Following the introduction, Chapter 2 presents the relevant theoretical background by reviewing the key concepts of traditional and behavioral finance, with a focus on herding behavior and market-wide herding measures. Chapter 3 provides a review of prior literature on herding in equity markets, prior findings on emerging equity markets, herding during extreme market conditions, and herding on up- and down-market days. Lastly, the study's hypotheses are presented based on prior literature. Chapter 4 introduces the data and descriptive statistics. Chapter 5 presents the methodology used in the empirical analysis. Chapter 6 presents the empirical results and discusses the study's limitations. Chapter 7 presents the conclusion of this thesis. Lastly, the appendix is provided at the end of the thesis.

ChatGPT (GPT-5.4 Thinking) was used in this thesis as a support tool to identify potentially relevant academic literature and refine academic language. Grammarly for Microsoft Word (version 1.2.258.1885) was used as a support tool for academic language refinement and grammar correction. The tools were not used as scientific sources. All academic literature cited in the thesis was independently evaluated and cited based on the original source material. The author remains fully responsible for the final content of the thesis.

2 Theoretical background

The theoretical foundations of finance can be broadly divided into traditional finance and behavioral finance (Baker & Ricciardi, 2014). Traditional finance is built on assumptions such as efficient markets, rational investor behavior, and the trade-off between risk and return. Theories, including the efficient market hypothesis and the capital asset pricing model, are cornerstones of fundamental beliefs, such as that asset prices reflect all available information and that deviations from fundamental values are only temporary. Empirical research, however, has shown that traditional finance theories have limitations, as investor behavior often differs from the expected behavior of a rational investor. Behavioral finance emerged to explain investor behavior more broadly. It argues that investment decisions are often influenced by different psychological and social factors. This chapter first reviews some of the main theories of traditional finance. It then discusses behavioral finance, with a focus on herding behavior. Lastly, this chapter reviews the widely used market-wide herding measures.

2.1 Traditional finance

Traditional finance views decision-making as a rational and systematic process. Baker and Ricciardi (2014) argue that the foundations of traditional finance are closely linked to the classical decision theory (CDT), which explains how individuals make choices between different possible courses of action. The theory assumes that people are able to rank their preferences across various possible outcomes and assign probabilities to each potential outcome. According to CDT, decision-making is a normative process that is aimed at identifying the optimal choice that maximizes expected benefits under uncertainty. Furthermore, it assumes that individuals act as rational agents that are driven by self-interest and seek to maximize their benefits under certain constraints. CDT is a fundamental theory in traditional finance, as it relies on consistent preferences, the ability to assess risk, and logical reasoning.

Financial decision-making often takes place under uncertainty, and investors may not know all the possible future outcomes or the probabilities of those outcomes. Although the world is largely uncertain, theoretical models typically focus on risk rather than uncertainty, because with risk, the probabilities of possible outcomes are known. However, even after accounting for risk, investors still face the question of how to identify the optimal decision (Baker & Ricciardi, 2014). To address this issue, Von Neumann and Morgenstern (1944) developed the expected utility theory (EUT) as a way to explain decision-making under risk. The theory assumes that individuals have rational preferences and they choose the option that gives them the highest expected utility. Therefore, EUT defines rational choices in risky environments. The theory is based on the idea of a utility function, where individuals can rank alternatives according to their preferences, and a rational person then chooses the alternative with the highest expected utility. The utility function shows that utility increases with wealth, but the increase becomes smaller at higher levels of wealth, as illustrated in Figure 1. It also captures the trade-off between risk and return, which is a central idea in traditional finance (Baker & Ricciardi, 2014). Investors are generally assumed to be risk-averse, meaning that when expected returns are equal, they prefer the option with less uncertainty. The utility function also provides a way to display risk aversion by showing how utility changes as wealth changes.

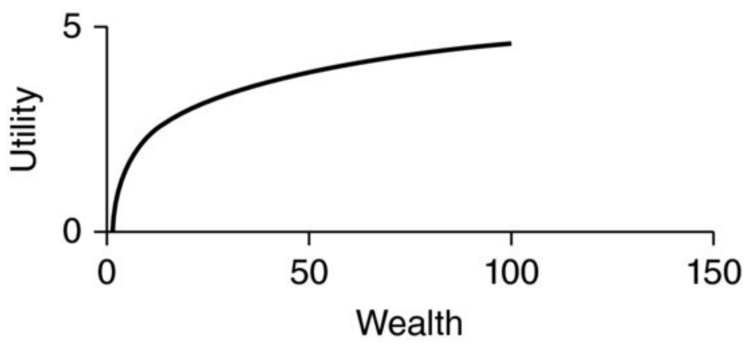

Figure 1. The Utility Function for a Risk Averter (Baker & Ricciardi, 2014, p. 25)

2.1.1 Efficient market hypothesis

The efficient market hypothesis (EMH) has been a prevailing concept in finance for several decades. According to the EMH, competitive markets integrate all available information into asset prices, which prevents investors from consistently achieving superior returns (Baker & Ricciardi, 2014). Fama (1970) further categorized the EMH into three different forms based on the level of information that is reflected in market prices. Weak-form efficiency refers to the situation in which prices reflect past information, meaning that technical analysis of historical price movements will not yield abnormal returns. Semi-strong efficiency means that prices reflect all publicly available information and past asset price data. It implies that investors cannot earn abnormal returns by analyzing earnings announcements or following news reports. The last form is the strong-form, where asset prices reflect all publicly available information, as well as any private information. This would mean that even companies' insiders would not be able to profit from abnormal returns. The author notes that the strong-form efficiency is not a realistic scenario and should primarily be considered a benchmark for assessing deviations from market efficiency. However, Fama (1970) finds empirical support for the weak-form efficiency.

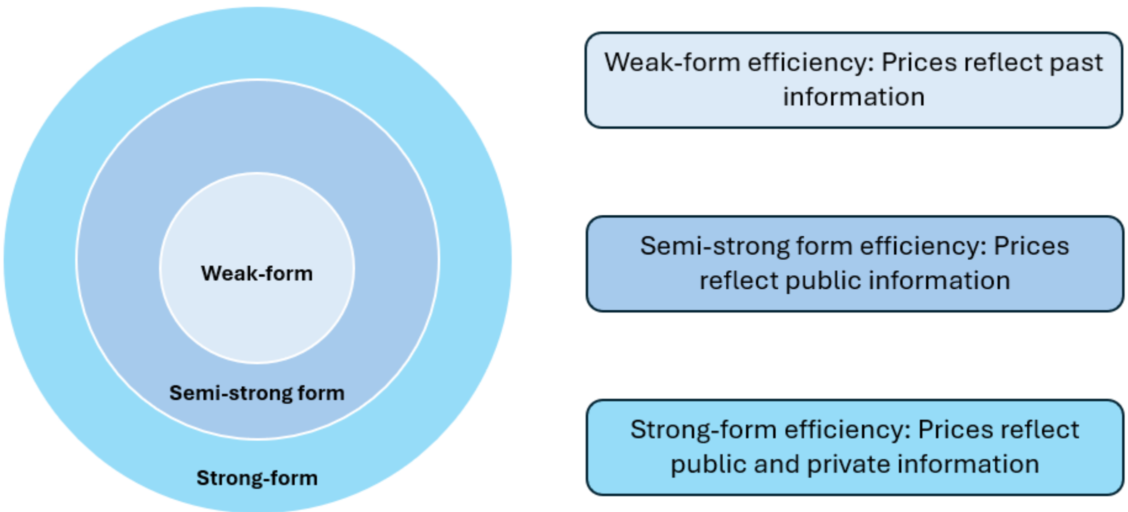

Figure 2. Efficient Market Hypothesis: Different Levels of Market Efficiency

Kendall and Hill’s (1953) findings on stock and commodity prices drew significant attention in the 1950s and are considered one of the first and most important evidence supporting the EMH. According to them, weekly price changes show little or no connection with price changes during the following week. This suggests that historical price data does not provide a reliable basis for predicting future price movements, which would imply that it is impossible to predict price movements solely based on past data. Therefore, asset prices would follow a random walk, making future price fluctuations unpredictable. The random walk theory is closely linked to the EMH, as it indicates that past changes in asset prices cannot be used to forecast future prices.

The EMH started to face criticism, and by the early 2000s, it no longer held the same level of dominance among academics that it once did (Malkiel, 2003). Financial economists and statisticians argued that stock prices show at least some degree of predictability. Scholars in behavioral and psychological finance highlighted the role of human behavior in price formation, suggesting that future prices can be at least partially anticipated using historical price trends and specific fundamental indicators. Some even advanced the more controversial view that such patterns could allow investors to achieve returns exceeding those expected on a risk-adjusted basis. The criticism concerned several areas, including evidence of predictability and anomalies, behavioral explanations of irrationality, and significant market events such as bubbles.

Empirical studies in the 1980s and 1990s revealed several return patterns that conflict with the EMH. Shiller (1981), for example, showed that stock prices change significantly more than can be explained by subsequent changes in dividends, suggesting that excess volatility exists, which is inconsistent with rational pricing. De Bondt and Thaler (1985) found evidence of investor overreaction, where assets that had performed poorly tended to outperform previous winners over time. Jegadeesh and Titman (1993) found momentum effects, where past winners continued to generate abnormal returns in the short-term, which challenges the idea that price changes are unpredictable. Fama and French (1992) also challenged the EMH by identifying size and value anomalies that traditional asset pricing models failed to explain. Together, these studies suggest that market returns are at least partly predictable.

Behavioral finance research provides further theoretical and empirical evidence that psychological biases drive market inefficiencies. Shleifer and Summers (1990) argued that noise traders and irrational investors may move prices away from their fundamental values, and that rational arbitrageurs may be unable to fully correct these mispricings. Barberis et al. (1998) later proposed a model of investor sentiment that shows how behavioral biases can create predictable patterns, such as momentum and reversals. Hirshleifer (2001) also reviewed evidence that cognitive errors and herd behavior systematically affect asset prices, which may lead markets to deviate from efficiency. Overall, these studies challenge the EMH and demonstrate that markets are influenced not only by rational expectations but also by persistent behavioral factors. While numerous studies have documented systematic deviations from the EMH, it still remains a foundational theory in finance, providing a useful benchmark for assessing market efficiency and for identifying inefficiencies that may arise from behavioral factors.

2.1.2 The Capital Asset Pricing Model and the Security Market Line

While the EMH provides insight into how information is incorporated into asset prices, it does not directly address how investors are compensated for bearing different types of risk. To address the relationship between risk and return, Sharpe (1964), Lintner (1965), and Mossin (1966) developed the capital asset pricing model (CAPM). Building on modern portfolio theory (MPT) introduced by Markowitz (1952), the CAPM explains the risk-return trade-off by introducing beta, which measures systematic risk. The model suggests that an asset’s expected return should be linearly related to its exposure to market risk, therefore providing a testable way for asset pricing. The CAPM is still considered one of the most important theories in modern finance.

As mentioned, the CAPM builds on the theory of MPT. Therefore, one needs to understand the key aspects of MPT. The theory is founded on the principle that investors are risk-averse and therefore demand compensation in the form of a risk premium for holding uncertain assets (Markowitz, 1952). Risk is commonly measured by the variability of returns, where a higher standard deviation indicates greater uncertainty. Markowitz's main contribution was to show that combining assets into a portfolio can lower overall risk, as long as the assets are not perfectly correlated. The risk of the portfolio depends on the correlation of assets' returns, and diversification reduces risk whenever securities' returns are less than perfectly positively correlated. According to Markowitz (1952), including a risk-free asset, typically a U.S. treasury bill, enables investors to determine which risky portfolio offers the most favourable risk–return trade-off. This portfolio is known as the optimal risky portfolio, also referred to as the market portfolio, and all investors should hold this same portfolio, adjusting only the proportion of wealth allocated to it based on their individual risk preferences. When there is no basis to assume that some investors would possess superior information, everyone should choose the market portfolio as the most efficient portfolio (Baker & Ricciardi, 2014). The key aspect of MPT is that diversification eliminates unsystematic (diversifiable) risk, leaving only systematic (market) risk that cannot be diversified away.

Since unsystematic risks can be diversified away, and investors are considered risk-averse, according to CAPM, investors are willing to bear more risks only if they are compensated with higher expected returns. The CAPM is based on some key assumptions. First, regarding investor behavior, investors are always assumed to be mean-variance optimizers with a similar time horizon and the same set of inputs and information. Second, markets are assumed to operate efficiently with minimal barriers to trading. Provided that unsystematic risk can be diversified away, then only systematic risk should matter for expected returns. Therefore, in CAPM, an asset's risk should be assessed by its contribution to the overall risk of a diversified portfolio or a market portfolio. To measure a security's systematic risk, the CAPM introduces beta (β). Beta measures how strongly an individual asset's return responds to movements in the overall market. Every asset has its own beta, while the market portfolio always has a beta of 1, and the risk-free asset has a beta of 0. A beta of one means the asset's return is expected to move in line with the market's return. A beta value greater than one indicates that the asset is more volatile than the market. Likewise, a beta less than one indicates that the asset is less volatile than the market. In practice, beta is estimated by taking the covariance between the asset's return and the market return, divided by the variance of the market return (Baker & Ricciardi, 2014; Bodie et al., 2023). By using the above information, the following equation of CAPM can be made:

$$E(R_i) = R_f + \beta_i\big(E(R_m) - R_f\big) \tag{1}$$

Where $E(R_i)$ the asset's expected return, R_f is the risk-free interest rate, $E(R_m)$ is the market's expected return, and β_i is the beta or risk of an asset. The expected market risk premium $\big(E(R_m) - R_f\big)$ should always be a positive value. If it were negative, a rational investor would never invest in a risky asset.

Assuming investors optimize their portfolios according to Markowitz's (1952) efficient portfolio diversification, they can construct an efficient frontier that shows all possible risky portfolios and identifies the most efficient one. The efficient frontier identifies the best available portfolio of assets before considering the risk-free asset. The capital market line (CML) represents the risk–return trade-off for efficient portfolios that combine the risk-free asset with the market portfolio (Bodie et al., 2023). For individual assets, the security market line (SML) shows the relationship between expected return and systematic risk, which is measured by beta.

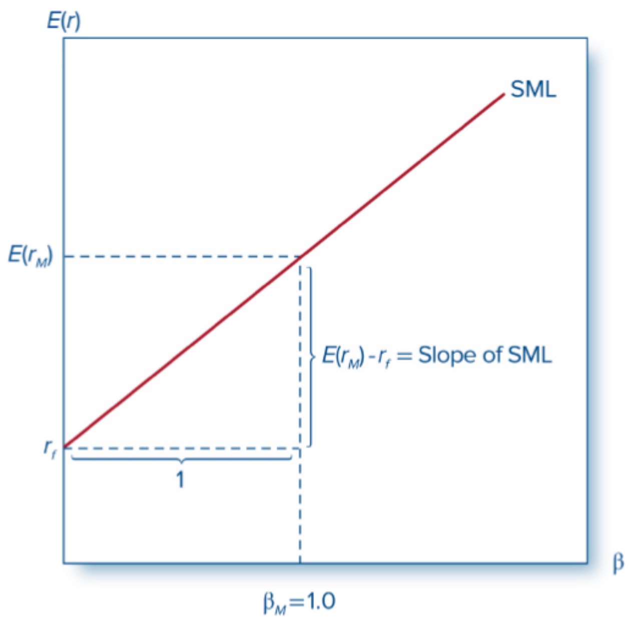

Figure 3. Security Market Line (Bodie et al., 2023, p. 291)

The difference between CML and SML is that CML demonstrates how an efficient portfolio's risk premium varies with the portfolio's standard deviation. In contrast, the SML illustrates how an individual asset's risk premium varies with its market risk.

Although the CAPM has played a central role in financial economics, it has been widely criticized both theoretically and empirically. As mentioned earlier, the theory builds on various assumptions, which Bodie et al. (2023) divide into two categories: individual behavior and market structure. First, investors are always assumed to be mean-variance optimizers, to have a single-period investment horizon, and to have the same input lists of assets. Second, all assets are available to everyone on the market, investors can borrow and lend at the risk-free rate, and there are no transaction costs or taxes. Several studies have identified flaws in these assumptions, arguing that many of them are unrealistic. For example, Roll (1977) argued that the CAPM is theoretically untestable because its central benchmark, the market portfolio, cannot be observed. The model requires a portfolio that includes all risky assets in the economy, which essentially means that all asset classes should be included. However, in practice, empirical tests often use only stock indices as proxies.

Empirical evidence regarding anomalies, such as size, value, and momentum, is also found to explain expected returns. Fama and French (1992) challenged the central prediction of the CAPM, that expected returns are determined solely by an asset's beta. They found that the relationship between beta and average returns is weak, whereas firm size and book-to-market equity ratio explained much of the cross-sectional variation in returns. Lakonishok et al. (1994) provided evidence that so-called value strategies consistently generate higher returns than would be expected given their beta, or systematic risk. The authors argue that the overperformance is related to systematic investor biases. Furthermore, Jegadeesh and Titman (1993) found that momentum strategies, which buy stocks that have performed well in the past and sell those that have performed poorly, generate significant positive abnormal returns. Their evidence suggests that momentum profits arise from delayed market reactions to firm-specific information, rather than from compensation for market risk. To conclude, many studies suggest that beta alone is not enough on its own to capture risk and that stock returns are better explained by multiple risk factors, leading to the development of multifactor models for asset pricing.

2.1.3 Multifactor models

Because the CAPM could not fully explain the variation in stock returns, researchers began to develop multifactor models. Ross (1976) introduced the arbitrage pricing theory (APT), which provides the theoretical basis behind multifactor asset-pricing models. The key point of APT is that a single factor is insufficient for determining market risk and that systematic risk comprises several factors. However, the APT does not specify which factors influence systematic risk. Roll and Ross (1980) provided the first empirical evidence, supporting the APT, that more than one common factor influences systematic risk. According to Bodie et al. (2023), a generalized equation of a multifactor model is the following:

$$R_i = E(R_i) + \beta_{i1}F_1 + \beta_{i2}F_2 + e_i \qquad (2)$$

Where R_i is the asset's excess return, $E(R_i)$ is the asset's expected excess return, β_{i1} and β_{i2} are the betas for risk factors F_1 and F_2 and e_i is the idiosyncratic or firm-specific risk. According to Bodie et al. (2023), the expected value of the factors is zero because they essentially quantify the surprise in the systematic variable.

The Fama–French Three-Factor Model is often considered as one of the most influential of the multifactor models. Its empirical basis started with Fama and French (1992), who found that the firm size and book-to-market value help explain cross-sectional variation in average stock returns. More specifically, smaller firms and firms with high book-to-market ratios tended to earn higher returns than predicted by beta alone. These results suggested that risks are multidimensional and cannot be completely captured by a single market factor. Fama and French (1993) also developed the three-factor model, which adds market capitalization (SMB) and book value (HML) factors alongside the market risk premium as systematic sources of asset returns. The equation for the Three-Factor Model is as follows:

$$R_{it} = \alpha_i + \beta_{i1} R_{Mt} + \beta_{i2}\,\mathrm{SMB_t} + \beta_{i3}\,\mathrm{HML_t} + e_{it} \quad (3)$$

Where, in addition to equation (2), α_i is the abnormal return or alpha, $\beta_{i1} R_{Mt}$ is the asset's sensitivity to the market factor, $\beta_{i2}\,\mathrm{SMB_t}$ is the asset's sensitivity to SMB (small minus big factor), and $\beta_{i3}\,\mathrm{HML_t}$ is the asset's sensitivity to the HML (high minus low) factor.

Fama and French (2015) later extended their model to a five-factor model by adding profitability (RMW) and investment (CMA) factors. Using U.S. data from 1963–2013, the five-factor model explains returns better than the three-factor model across portfolios sorted by size, value, profitability, and investment. Furthermore, they found that once RMW and CMA are included, HML (value) contributes little on average because its premium is largely absorbed by exposures to the new factors. They note that the main weakness is the model's poor performance for small firms that invest aggressively despite weak profitability. The equation of the five-factor model is as follows:

$$R_{it} = \alpha_i + \beta_{i1}R_{Mt} + \beta_{i2}SMB_t + \beta_{i3}HML_t + \beta_{i4}RMW_t + \beta_{i5}CMA_t + e_{it} \quad (4)$$

Where, in addition to equations (2) and (3), $\beta_{i4}RMW_t$ is the asset's sensitivity to RMB (robust minus weak) factor and $\beta_{i5}CMA_t$ is the asset's sensitivity to CMA (conservative minus aggressive) factor.

There has been extensive research seeking factors that could explain security returns beyond the three-factor model. Fama and French (2018) go even further, developing a six-factor model that introduces momentum as an additional factor. However, they note that without a solid theory to determine which factors to use or add, researchers can keep adding variables until the model's factor mix effectively replicates the efficient mean-variance (tangency) portfolio for that specific sample period. They emphasize that this is ''data dredging'' meaning that it fits past returns but doesn't say much about true, generalizable risk. Nevertheless, the important thing to understand about multifactor models is that several factors likely influence the expected return-risk relationship of assets, contrary to what the CAPM suggests.

2.2 Behavioral finance

Behavioral finance emerged as a response to the limitations of traditional theories such as the EMH and the CAPM, which assume rational investors and fully efficient markets. While traditional finance focuses on how investors should make decisions, behavioral finance focuses more on why investors actually make the decisions they do. Behavioral finance gained more attention in the 1980s and 1990s, as empirical findings increasingly challenged traditional finance theories, and major market events, such as the 1987 stock market crash and the 1990s internet stock bubble, drew attention to the importance and role of investor behavior (Baker & Ricciardi, 2014). Behavioral finance is strongly influenced by psychology, especially Kahneman and Tversky's (1979) prospect theory, which showed that decision-making often differs from the rational behavior that is assumed in EUT. As evidence of investor irrationality increased, it further decreased the status of traditional finance.

As noted earlier, EUT explains decision-making under risk by assuming that individuals have rational preferences and that they prefer more to less. It provides a normative framework of how people should behave when trying to maximize their utility in uncertain situations. Kahneman and Tversky (1979) introduced the prospect theory as an alternative to EUT, offering a model of how people actually make choices under risk. While both theories share similarities, prospect theory proposes that individuals maximize perceived value rather than expected utility. It also suggests that people evaluate outcomes relative to a reference point, focusing more on changes in wealth rather than final states. According to the authors, individuals tend to be risk-averse when facing potential gains but risk-seeking when facing potential losses. Moreover, the psychological impact of losses is generally greater than that of similar gains, as evidenced by the fact that losing 500 in wealth causes more pain than the pleasure from a 500 increase in wealth, as illustrated in Figure 4.

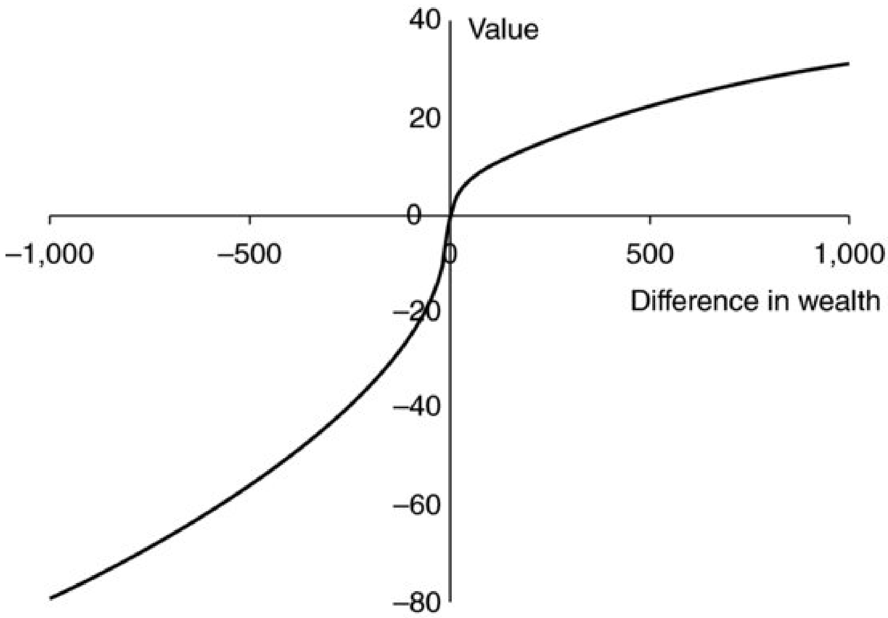

Figure 4. The Prospect Theory Value Function for a Loss Averter (Baker & Ricciardi, 2014, p. 30)

According to Ritter (2003), behavioral finance can be divided into two parts: cognitive psychology and limits to arbitrage. Cognitive psychology is related to how individuals think. Psychological research shows that people often make systematic errors in their thinking. Cognitive psychology has identified numerous systematic patterns in human behavior, and in behavioral finance literature, these are often referred to as behavioral biases. Without delving into the various biases, the relevant point is that biases such as overconfidence and herding can affect investor decision-making. According to Hirshleifer (2015), many of the psychological biases examined in behavioral finance can be explained by heuristic simplification, self-deception, and affective short-circuiting. These factors also shape the processes of psychological updating that sustain behavioral biases, even in situations when individuals could learn from their past mistakes.

Limits to arbitrage are another key concept in behavioral finance because they help to explain why market inefficiencies and mispricings may persist for long periods of time, even when rational investors are present in the market. Arbitrage means trading to profit from price differences across markets, but it can either improve or worsen the price efficiency of an asset. When investors disagree, prices reflect a mix of beliefs, so both irrational traders and rational arbitrageurs may see themselves as exploiting mispricing (Hirshleifer, 2015). According to Shleifer and Vishny (1997), arbitrage should eliminate mispricing by allowing traders to sell overpriced assets and buy underpriced ones until prices return to their fundamental values. In practice, however, arbitrage is risky, costly, and often constrained. Short-sale restrictions, limited capital, and uncertainty about the timing of price corrections can prevent arbitrageurs from correcting mispricings. Furthermore, arbitrage tends to be weakest when mispricing is greatest, which suggests that rational investors may avoid markets with severe distortions. For this reason, limits to arbitrage are regarded as one of the two foundational pillars of behavioral finance, alongside cognitive psychology (Ritter, 2003; Hirshleifer, 2015)

2.2.1 Theoretical background of herding behavior

Research on herding gained momentum in the 1990s, after the growing acceptance of behavioral finance among researchers. The concept of herding appears across a wide range of research fields and contexts. In finance and economics, herding behavior refers to the tendency of economic agents to follow one another’s actions and disregard their own information in financial markets (Spyrou, 2013). In these circumstances, investors tend to imitate others’ actions and use that as a reason for making their investment decisions. Theoretical models explaining herding behavior in financial markets are generally grouped into two broad categories: those based on rational decision-making and those associated with non-rational behavior. However, there exists a gap in the literature in models that integrate both sources of herding. Furthermore, many existing models yield implications that are difficult to test empirically, given the limitations of the available data.

The origins of herding research are often linked to Keynes (1936). He emphasized that collective behavior and psychological factors influence the financial markets. While he did not define herding as such, his paper influenced later thinking about crowd-driven market behavior. Banerjee (1992) and Bikhchandani et al. (1992) developed some of the earliest herding models in the 1990s, establishing herding as a rational, information-based phenomenon. Devenow and Welch (1996) added further insight to this view by arguing that herding can be rational at the individual level, but at a broader social level, it is inefficient. Later behavioral finance researchers, such as Shiller (2000), emphasized the role of psychological biases and investor sentiment as drivers of herding and market anomalies, therefore highlighting the irrational part of herding.

In the finance literature, researchers have proposed several different definitions for herding. Bikhchandani et al. (1992) define herding as informational cascades, arguing that individuals observe others' actions and rationally choose to mimic them because the behavior of predecessors is viewed as superior to their own information. Devenow and Welch (1996) describe herding as correlated actions among individuals that may result from the same information arriving simultaneously at different individuals. According to them, herding is either the result of some coordinated information, a price movement, for example, or the ability of investors to follow others' actions. Herding is seen as a phenomenon arising from either rational or irrational motives. The authors argue that the rational reasons stem from compensation-based herding, reputational herding based on career concerns, and informational herding. Scharfstein and Stein (1990) define herding from an institutional perspective. They argue that under uncertainty, professional managers may copy their peers' investment decisions even when they have their own private information. Although this behavior may be inefficient, it may still be rational if the manager is concerned about their own reputation in the labor market. Shiller (2000) defines herding more broadly as a process in which people follow one another based on narratives, psychological reasons, and social influence. Therefore, he emphasizes the behavioral perspective on herding and argues that it often occurs without full consideration of fundamental factors. Many more definitions of herding exist, but the key point is that it is a complex phenomenon influenced by both rational and behavioral factors.

2.2.2 Herding dimensions

Herding can be categorized into different dimensions in two main ways. The first categorization is based on what drives herding behavior, and the second on whether herding is rational or irrational. Bikhchandani and Sharma (2001) distinguish between spurious and intentional herding. Spurious herding occurs when investors make similar decisions because they are affected by the same information or they respond to changes in fundamental factors. For example, a significant increase in interest rates may lead many investors to sell equities simultaneously because higher rates may negatively impact their future value as investments. In this case, the behavior is considered rational market behavior rather than intentional herd behavior. Spurious herding is often regarded as rational and efficient, since it reflects investors' responses to changes in available information. While spurious herding may appear to involve herding in market data, it is correlated trading based on common information, shared constraints, or mutual responses to fundamentals, rather than imitation of other investors.

Intentional herding refers to investors intentionally copying one another for various reasons (Bikhchandani & Sharma, 2000). It is often seen as a major cause for excessive volatility, systemic risks, and fragile markets, or in other words, inefficient outcomes. The reasons for Intentional herding can be both rational and irrational. Bikhchandani and Sharma identify three key factors underlying rational herding: reputational concerns, compensation structures, and imperfect information. While these factors may sometimes lead to inefficient market outcomes, they may be rational from the individual’s perspective. For example, it may be rational for a fund manager to act as a herd to safeguard his/her own reputation. Irrational herding is typically associated with psychological and social factors that drive investor herding. Irrational herding may emerge from various psychological stimuli and different restraints, such as investors being influenced by social factors to copy each other under uncertain periods (Keynes, 1936), information scarcity and the use of common heuristic rules (Baddeley et al., 2004), and irrational noise traders who are impacted by systematic biases (Black, 1986). However, empirically distinguishing between spurious and intentional herding, as well as between rational and irrational herding, is difficult. Investment decisions are typically influenced by multiple factors, and existing empirical methods primarily test for clustering and co-movements, which cannot clearly separate between the different dimensions of herding (Bikhchandani & Sharma, 2001).

2.2.3 Root causes of herding

Herding in financial markets is a complex phenomenon that likely results from multiple underlying causes. Several studies have tried to identify its causes, and many potential explanations have been proposed. Banerjee (1992) provides an information-based explanation by introducing the concept of informational cascades as one potential factor causing herding behavior. The author argues that herding typically occurs when information is uncertain, as individuals may then rely more on the actions of others or on publicly accepted opinions. Banerjee emphasized that people observe the choices of those who act before them and make their own decisions accordingly, because others' choices may reveal private information that they possess. Therefore, it is rational for later individuals to follow those who have acted before them, even if the individual's own information would suggest otherwise. However, once enough people disregard their own information and follow others, the following decisions become uninformative. This leads to an informational cascade, which creates herd behavior where individuals disregard their own private information. Bikhchandani et al (1992) build on Banerjee's concept of informational cascades and argue that these cascades occur when it becomes optimal for individuals to follow others and ignore their own information. According to them, once cascades form, decisions are no longer based on private information, which can lead individuals to act on incorrect information. The main point of their model is that cascades are fragile: they rely on little actual information, and therefore even small shocks or new information can change the crowd's behavior. The fragility helps explain why herding can often lead to extreme market events and crashes.

Reputational and career concerns are another factor that can lead to herding behavior. Scharfstein and Stein (1990) argue that managers may, under certain conditions, replicate their peers' investment choices while ignoring valuable private information. They explain reputational herding through a "sharing-the-blame" mechanism. Since skilled managers are assumed to observe similar signals about fundamental information, imitating peers' decision signals to the labor market that the manager's information is consistent with others', and that the manager is therefore more likely to be competent. In contrast, taking a contrarian position may make the manager look less informed or less competent. Although this can be inefficient from a broader societal perspective, it can be rational for managers who prioritize protecting their reputations in the labor market. Devenow and Welch (1996) also describe the previously mentioned behavior as "rational herding," since individuals act rationally given their reputational incentives. Many empirical studies support this view, including Graham (1999) and Hong et al. (2000), who provide evidence of analysts' reputational herding. For example, Graham finds that analysts with higher prior reputation and lower abilities (meaning less precise private information) are more likely to herd. The author demonstrates that analysts align their recommendations with the consensus to protect their reputations and overlook their own information.

A short investment horizon for the investor may also contribute to herding (Froot et al., 1992). Speculative investors with a short-term investment horizon might herd on the same information and attempt to benefit from investors with superior knowledge, rather than using a diversified set of information sources. As more speculators acquire a specific signal and trade based on it, that signal becomes increasingly embedded in prices. This makes early acquisition of the same information profitable and, thus, also potentially rational from a short-term investor's perspective. However, it may lead to inefficient outcomes if the information is of poor quality or unrelated to the asset’s fundamentals.

Noise trading, introduced by Black (1986), is another potential factor contributing to herd-like behavior in markets. Black introduced noise trading as a significant but irrational form of correlated trading based on "noise" rather than on actual useful information about an asset. Noise traders’ behavior is prone to behavioral biases, and thus, they can reduce the information content of asset prices. The author offers two possible explanations for why investors trade on noise: either they interpret noise as information, or they simply prefer to trade even if they would be better off not trading. While Black does not argue that noise trading creates herding, it can increase coordinated sentiment-driven trading that is not based on fundamentals and therefore move prices in ways that are not easily arbitraged away. This can increase herding in financial markets.

Several psychological and social factors can generate herd-like behavior in financial markets. Shiller (2000) argues that people have psychological anchors for the market, specifically quantitative and moral anchors. A quantitative anchor is a simple numerical reference point, such as a recent price level, especially when investors lack a clear view of the market's fundamental value. A moral anchor is a reason to believe that it is acceptable to invest or remain invested (e.g., a narrative) and is used under uncertain conditions. If these anchors become commonly held beliefs among investors, they may serve as shared reference points between different investors. This may result in a higher correlation among investors' beliefs and increase the probability of crowd-following, especially when fundamental valuation is uncertain. Although Shiller (2000) does not directly claim that psychological anchors cause herding, they may still contribute to it. Furthermore, people who communicate regularly often begin to think in similar ways, and market-related beliefs can spread quickly through word-of-mouth and broader communication channels, especially during times of market crises. This may create correlated beliefs among investors, which can lead to speculative bubbles and herding. Baker and Ricciardi (2014) also emphasize investors' uncertainty about future prices and their limited ability to process complex information as potential causes of herding, which may encourage them to rely on the "wisdom of crowds" or to follow perceived market leaders, such as investment professionals. Consistent with Keynes's (1936) beauty-contest logic, investors tend to try to predict and match what they expect others will believe and do. Especially when fundamental value is difficult to determine, following the crowd may seem rational and may, therefore, serve as a potential cause of herding.

2.3 Market-wide herding measures

Measures for empirically studying herding can be divided into two categories. The first consists of methods that use proprietary or micro-level data, which examine the herding behavior of specific investor types. The most widely used methods in that group include those of Lakonishok et al. (1992) and Sias (2004). The second, which is the focus of this thesis, is based on market activity or aggregate price data. These studies examine herding toward market consensus and are typically referred to as market-wide herding measures (Spyrou, 2013). Several market-wide herding measures have been proposed, but this chapter will focus on the widely adopted ones that use cross-sectional stock return data. This section introduces the first widely used market-wide herding measures by Christie and Huang (1995) and Chang et al (2000), as well as the modified method by Chiang and Zheng (2010), which is used in this thesis.

Christie and Huang (1995) tested market-wide herding by examining clustering of individual stock returns around the market return during periods of market volatility. The authors argued that during extreme market movements, investors tend to ignore their own information and trade in line with market consensus. Therefore, individual asset returns should not deviate significantly from the market return, resulting in relatively low dispersion. This contradicts rational asset-pricing models, which predict that periods of high volatility are associated with higher dispersion. However, according to the authors, market participants' investment decisions are influenced by the overall market state. During periods of significant market movements, investors follow the collective actions of the market, leading individual stock returns to cluster around the overall market return. The authors estimate the cross-sectional standard deviation (CSSD) of returns as an indicator of herd behavior with the following formula:

$$CSSD_t = \sqrt{\frac{\sum_{i=1}^{N}\left(R_{i,t}-R_{m,t}\right)^2}{N-1}} \tag{5}$$

Where $R_{i,t}$ is the observed return on firm *i* at time t, $R_{m,t}$ is the cross-sectional average return at time *t*, and N is the number of stocks in the market portfolio. The CSSD metric measures the cross-sectional dispersion of individual asset returns around the average portfolio return. If herding exists, cross-sectional dispersion is expected to be low because returns cluster around the portfolio return. However, low dispersion alone is not sufficient evidence of herding because, for instance, little or no information arriving during a trading period can cause low return dispersion even when investors are not herding at all. Therefore, Christie and Huang (1995) employ the following regression to investigate whether return dispersions significantly decrease during periods of pronounced market movements:

$$CSSD_t = \alpha + \beta_1 D_t^L + \beta_2 D_t^U + \epsilon_t \tag{6}$$

Where D_t^L is a dummy variable equal to 1 when the market return on day t lies in the extreme lower tail of its distribution, and 0 otherwise. Similarly, D_t^U equals 1 when the market return on day t lies in the extreme upper tail of the distribution, and 0 otherwise. The intercept α represents the average return dispersion for all periods outside the extreme-market regions captured by the two dummy variables. Using the dummy variable approach, it is possible to distinguish investor behavior under highly positive and highly negative market conditions from that observed under normal market conditions. Christie and Huang (1995) argue that under rational asset-pricing models, the coefficients β_1 and β_2 are expected to be significantly positive. Negative and statistically significant coefficients for β_1 and β_2 would be consistent with herd behavior.

Because CSSD is based on squared deviations, it is relatively sensitive to extreme returns. Chang et al. (2000) introduce a closely related market-wide approach using the cross-sectional absolute deviation (CSAD), which reduces the influence of outliers and allows for a nonlinear test for herding. Their key argument is that if investors increasingly follow the overall market during periods of significant average price movements, the linear, increasing correlation between dispersion and market return may break down. It could therefore become nonlinearly increasing or even decreasing as market movements increase. To capture this possibility, they model the relationship between the cross-sectional absolute deviation (CSAD) of returns and the market return using a nonlinear regression specification. To define the expected cross-sectional absolute deviation of stock returns (ECSAD), they initiate with the conditional version of the Black (1972) CAPM in the following way:

$$ECSAD_t = \frac{1}{N}\sum_{i=1}^{N}|\beta_i - \beta_m|E_t(R_m - \gamma_0) \qquad (7)$$

Where R_m denotes the return on the market portfolio, γ_0 is the return on the zero-beta portfolio, β_m represents the systematic risk (beta) of the equally weighted market portfolio, and β_i represents the systematic risk (beta) of any asset *i*. Chang et al. (2000)

subsequently illustrate that the increasing and linear correlation between dispersion and time-varying market expected returns can be shown with the following equations:

$$\frac{\partial\, ECSAD_t}{\partial\, E_t(R_m)} = \frac{1}{N}\sum_{i=1}^{N}|\beta_i - \beta_m| > 0, \qquad (8)$$

$$\frac{\partial^2\, ECSAD_t}{\partial\, E_t(R_m)^2} = 0. \qquad (9)$$

These equations express the rational asset-pricing prediction that expected return dispersion $ECSAD_t$ increases linearly with the expected market return. Building on equations (8) and (9), Chang et al. (2000) propose a herding test that includes a term to allow for nonlinearity in the relationship between asset return dispersion and the market return. To approximate the unobservable $ECSAD_t$ the authors utilize the CSAD at time t as its empirical proxy. CSAD is calculated by averaging the absolute deviations of each stock's return relative to the equally weighted market portfolio's return. The rationale for this methodology is that under rational asset pricing, dispersion should increase with the absolute market return. However, if herding emerges during extreme market conditions, CSAD will increase less than proportionately and may even decline. Herding is identified by examining the relationship between the market return and CSAD. It is crucial to remember that CSAD is not intended to be used directly as a herding measure, but herding is reflected in the model's parameter estimates. The quadratic regression equation for capturing the relationship between CSAD and market portfolio return is the following:

$$CSAD_t = \alpha + \gamma_1 R_{m,t} + \gamma_2 R_{m,t}^2 + \varepsilon_t \qquad (10)$$

Where a negative γ_2 coefficient indicates the presence of herding behavior, and $R_{m,t}^2$ allows for a possible nonlinear relationship between the market return and CSAD. Chiang and Zheng (2010) introduce a modification to the regression specification (10) of Chang et al. (2000), offering an alternative approach to modeling asymmetry in investor behavior across different market conditions. The regression equation is as follows:

$$CSAD_t = \gamma_0 + \gamma_1 R_{m,t} + \gamma_2 \left|R_{m,t}\right| + \gamma_3 R_{m,t}^2 + \varepsilon_t \qquad (11)$$

Where in addition to equation (10), Chiang and Zheng (2010) include an additional signed market return $R_{m,t}$ to the equation, which allows for the inclusion of asymmetric investor behavior in different market directions by modeling both positive and negative conditions. Given its extensive use in the empirical herding literature and its suitability for testing market state-dependent and asymmetric herding, this approach provides a valid market-wide herding measure for this thesis.

3 Literature review

This chapter reviews the relevant literature regarding market-wide herding in equity markets and provides the foundation for the empirical analysis and hypotheses of this thesis. It first provides an overview of research on herding in financial markets. The chapter then reviews evidence of market-wide herding in emerging equity markets with a particular attention to India and China. Market characteristics of India and China are also reviewed. After this, it examines herding during extreme market conditions as well as during periods of up- and down-markets. The final section introduces the hypotheses of this thesis that are tested in the empirical part.

3.1 Overview of herding research in financial markets

Herding is a widely studied topic in financial research, and interest in the subject has increased substantially over time, especially after the 2008 financial crisis. According to Choijil et al. (2022), herding research can be broadly categorized into five sub-areas. The first focuses on a deeper understanding of herding behavior, while the second examines empirical results on herding and explores possible causes for it. The third concentrates on herding during crisis periods, the fourth investigates differences among investor types, while the fifth examines herding in portfolio management and its potential effects. This broad division of the literature highlights that herding is a complex and multidimensional phenomenon. This complexity may help explain why the herding literature has not found a common view of the phenomenon. In line with this, Choijil et al. (2022) note that despite the growth of herding research, there is still no consensus on the causes of herding.

Although herding research has advanced substantially, several limitations and open issues remain. According to Spyrou (2013), there is a gap among theoretical models, available empirical methodologies, and existing datasets. Many theoretical predictions, especially those related to informational cascades and investors' private information, are challenging to examine because the underlying decision-making processes are not observable in standard databases. For example, it is difficult to determine whether investors truly ignore their own information and imitate others' actions unless the researcher has access to the information available to them before the trade. As a result, empirical methods typically focus on indirect herding based on trading patterns, return comovement, or clustering of investment decisions. However, such clustering may arise for several reasons, which makes the identification of herding difficult. Different herding methods and limited access to the right kind of data may also explain some of the inconsistencies in herding research.

Another issue is that prior empirical studies have not produced fully consistent results (Spyrou, 2013). Regarding institutional investor herding, for example, some papers find little or no evidence of herding (Lakonishok et al., 1992; Grinblatt et al., 1995). On the other hand, Choi and Sias (2009) find evidence suggesting herding in US institutions, while Sias (2004) finds that institutions may herd by drawing conclusions from other institutions' trading behavior. Similarly, evidence regarding herding towards the market consensus in different markets varies, as, for example, some studies report herding in the Southern European markets (Caparrelli et al., 2004; Economou et al., 2011), while Chang et al. (2000) discover no evidence in Hong Kong and the US, some evidence in Japan, and no evidence in South Korea and Taiwan. More inconsistencies in empirical evidence exist, for example, regarding analyst herding and herding's impact on asset prices (Spyrou, 2013). Overall, it is reasonable to conclude that results in herding research may vary across markets and investor groups.

Chang et al. (2000) argue for the importance of studying herding across developing and more developed markets, as different market structures may affect how easily herding emerges in a given market. Such market features include structures such as the relative importance of institutional or individual investors, the quality of information disclosure, and the market's financial sophistication, for example, in the derivatives market. Furthermore, Chiang & Zheng (2010) note that empirical evidence from a select group of countries only shows how the local market behaves, and that broader implications for global herding cannot necessarily be drawn from those results. As herding may vary across markets, it is therefore reasonable to argue that herding should be examined in different market settings.

3.2 Herding in emerging equity markets

Prior literature often suggests that herding is more likely to arise in emerging equity markets than in advanced markets. Some researchers argue that one potential reason for this is related to greater information asymmetry in emerging markets (Chiang & Zheng, 2010). However, empirical evidence is mixed across different countries, market conditions, and sample periods. Chang et al. (2000) examine herding behavior across international markets and find evidence of herding in South Korea and Taiwan, but not in the US, Japan, or Hong Kong. They characterize South Korea and Taiwan as the only emerging markets in their study, suggesting that herding may be more prevalent in emerging markets. Chiang and Zheng (2010) study herding across 18 stock markets, providing a broad international sample. However, their results are more mixed regarding emerging markets. They do find herding in advanced and some Asian markets, but not in the Latin American markets. This suggests that herding is not necessarily always present in less-developed emerging markets and supports the argument that herding should be treated more as a market-specific phenomenon.

The findings of Chang et al. (2000), who document herding in South Korea and Taiwan, together with evidence from Demirer et al. (2010), who document herding in Taiwan, suggest that herding is not limited to the least-developed emerging economies. Those results rather indicate that herding may also be present in relatively sophisticated emerging markets. This strengthens the motivation to study herding in India and China, which can be considered relatively advanced emerging markets.

3.2.1 Characteristics of India and China as emerging markets

Some papers characterize the Indian and Chinese stock markets as relatively inefficient and riskier than more developed markets. The reasons are associated with factors including lower levels of investor sophistication, cultural differences, incomplete securities legislation, weak legal enforcement, less rigorous accounting and reporting practices, and limited transparency in information disclosure (Akbar & Samii, 2005; Khanna & Palepu, 2000; Li, 2008; Zhang & Zhao, 2004). Within the emerging market literature, some papers suggest that these shortcomings may be linked to herding behavior and market anomalies (Chang et al., 2000; Demirer & Kutan, 2006). However, given that much of this literature dates to the early 2000s and the Indian and Chinese economies have grown significantly since then, those assessments may not necessarily fully reflect current market conditions in these markets.

China and India both have several structural characteristics that support sustained long-term economic growth and strengthen their potential to play an increasingly important role in the global economy. These include at least a large geographic scale, an extensive labor force, a growing middle class, and supportive government policies (Lao & Singh, 2011). As their significance in the global economy has grown, it has become increasingly important for investors and researchers to understand the dynamics and the possibility of herding in their stock markets.

Compared with China, Lao and Singh (2011) suggest that the Indian equity market has been more strongly influenced by large financial institutions, which may contribute to more rational security analysis and less speculative trading. The Chinese stock market, by contrast, consists of A- and B-share markets, which differ substantially in investor composition. The A-share market has historically been dominated by domestic individual investors, who are often assumed to have less professional expertise and a more limited ability to access and interpret information accurately (Tan et al., 2008; Demirer et al., 2010). The B-share market has been dominated by foreign institutional investors (Tan et al., 2008). As individual investors are often assumed to be more prone to biases such as herding, herding might differ across these two market segments. However, this thesis uses the CSI 300 index, which comprises only A-shares. The analysis, therefore, reflects the behavior of the investors in the Chinese A-share market rather than both segments. The following two subsections will examine earlier findings on market-wide herding in each market.

3.2.2 Herding in Indian equity markets

Bhaduri and Mahapatra (2013) provide evidence of market-wide herding during market crashes in the Indian market. Using an alternative dispersion-based approach, they find that during extreme market movements, asset dispersion tends to decline rather than increase, consistent with herding. They also note that the extent of herding is time-varying and changes across years. During major crash periods, herding in the Indian equity market is more pronounced. Furthermore, the authors find that the increase in return dispersion during rising markets is relatively weaker than during declining markets, suggesting that herding may be stronger on up-market days.

Lao and Singh (2011) also find evidence of market-wide herding in the Indian stock market over the period 1999-2009. Their results further show that herding is conditional rather than constant in the Indian market, being more pronounced during periods of large market movements. They find results similar to those of Bhaduri and Mahapatra (2013), as herding appears to increase during upward market trends. Furthermore, herding in India seems unrelated to trading volume and is mainly present in mid-sized shares. Overall, their findings indicate that herding is not uniform in India and is weaker than in China. The authors associate the weaker herding mainly with the strong presence of financial institutions in India. Both Lao and Singh (2011) and Bhaduri and Mahapatra (2013) suggest that herding in India is more likely to arise under extreme market conditions. This motivates the present thesis, since the COVID-19 period constitutes a major market crisis in which herding behavior may possibly be observed.

3.2.3 Herding in Chinese equity markets

There exists more previous literature on herding in the Chinese market than in India. However, previous empirical studies have yielded mixed results. Demirer and Kutan (2006) find no evidence of herding in the Chinese equity market over the period 1999-2002. They find that dispersion in equity returns increases during large market movements, which suggests market behavior consistent with the predictions of rational asset pricing models in the Chinese market. However, the dispersion is much smaller during extreme declines, suggesting that stock returns behave more similarly in down markets. Unlike Demirer and Kutan, Tan et al. (2008) find significant herding in both A-share and B-share markets over the period 1997-2003, suggesting that herding is present in both market segments. Their results further indicate that herding occurs in both rising and falling markets, but is more pronounced in the A-share Shanghai market during periods of market rise. The B-share market shows no asymmetry.

Later studies generally provide stronger support for herding in China and show that herding is conditional on the market conditions and investor structure. Chiang et al. (2010) find evidence in both the Shanghai and Shenzhen A-share markets and show that investors herd in both rising and falling markets. In their main least squares method, however, they do not find comparable herding in B-shares, and B-shares show no herding during up-markets. These findings suggest that herding in China is more pronounced in the A-share segment, which is closely associated with domestic investors, rather than in the B-share segment, which is typically associated with foreign/institutional investors. Lao and Singh (2011) also find herding in China, and their results indicate that it is greater during declining markets, high trading volume, and large market volatility. Their findings further indicate that herding in China is conditional rather than constant. Liu et al. (2023) employ a group-specific method for herding and find that most-informed investors herd less than least-informed investors, though the difference is smaller during periods of uncertainty and market crises. Overall, prior herding literature from China indicates that herding is conditional and varies across market conditions and investor groups, highlighting the importance of studying herding during crises such as COVID-19.

3.3 Herding during extreme market conditions

Herding literature often assumes that herding should be more prevalent during extreme market conditions. Christie and Huang (1995), for example, argue that because individuals are more likely to ignore their own beliefs in favor of market consensus under unusual market conditions, herding is most likely to emerge during periods of market stress and large market movements. One possible explanation for this may be that herding becomes more likely when uncertainty is high, because investors treat others' observed actions as information rather than relying on their own private information (Bikhchandani & Sharma, 2000). However, empirical evidence on herding under extreme market conditions is mixed, with some studies providing evidence against it. This section presents the inconclusive literature on herding under extreme market conditions and provides the motivation for the subperiod analysis of the COVID-19 period in this thesis.

Christie and Huang (1995) find evidence inconsistent with herding during extreme market movements, meaning that dispersions increased during periods of market stress, which is more consistent with rational asset pricing models. Demirer and Kutan (2006) find similar evidence from the Chinese market, with no herding and rising return dispersion during large market movements, again supporting market efficiency and rational asset pricing models. Hwang and Salmon (2004) provide evidence from the US and South Korean markets that suggests that the Asian and Russian crises act as turning points that reduce herding rather than increase it. Their findings suggest that investors may turn to fundamental information and efficiency rather than aggregate market movements during crises. Chang et al. (2000) provide cross-country evidence from both developed and emerging equity markets. During extreme market movements, they find evidence of herding in the emerging markets of South Korea and Japan, but not in the developed markets of the US, Hong Kong, and Japan, suggesting that herding during large market moves is country-specific.

Several studies also find that herding becomes more apparent during extreme market environments and large market movements. Lao and Singh (2011) provide evidence that herding is more pronounced during large market movements in the Indian and Chinese equity markets. Moreover, the pattern is asymmetric between the two countries, since herding is stronger in extreme downward markets in China, whereas in India it occurs only during extreme up-markets. They also report that the 2008 financial crisis increased herding in China, whereas no herding was observed in India during that period. Their results suggest that herding during extreme market movements is not uniform and depends on market structure and direction. Similarly, Chiang and Zheng (2010) find evidence that herding is market-state-dependent. In their cross-country analysis, they show that herding is present in the US and Latin American markets during crisis periods, whereas full-sample herding is absent in these markets. Furthermore, they find that crisis periods trigger herding in the original crisis country and spread it to neighboring countries. Their evidence suggests that herding may become more prevalent during periods of crisis. Mobarek et al. (2014) provide further evidence from the European markets that herding depends on market conditions. While they find herding insignificant over the full sample period, they find that herding is particularly pronounced in most continental European markets during the global financial crisis and in Nordic markets during the Eurozone crisis. The findings indicate that herding may be weak during normal periods but may become pronounced during market crises. Overall, prior empirical evidence suggests that herding may become more pronounced under extreme market conditions, but the evidence remains mixed across countries and sample periods.

3.4 Herding during up- and down-markets

Examining herding over the full sample period, as well as during extreme market conditions, does not reveal whether herding differs due to market direction. Therefore, it is relevant to study herding separately in up- and down-markets, as prior literature suggests that investor herding may vary between rising and falling market periods.

Typical of the herding literature, the evidence on rising and falling markets is inconclusive. However, it supports the general conclusion that herding is asymmetric and may vary between the two directions, and is market-specific rather than universal.

Chang et al. (2000) provide evidence across five markets, showing that return dispersion is higher in rising markets than in down markets, suggesting that stocks move more similarly in down markets. Although their evidence regarding the directional asymmetry does not imply herding, their results are consistent with the idea that stocks tend to react sharply to negative news and more slowly to positive news. In another cross-country examination, Chiang and Zheng (2010) find that herding is present during both rising and falling markets in most advanced and Asian markets. However, the authors note that herding asymmetry in Asian markets is stronger, particularly during market uptrends, indicating that the direction of asymmetry is not uniform across studies. As for the US and South Korean markets, Hwang and Salmon (2004) provide evidence that herding is present in both bull and bear markets.

For India, Lao and Singh (2011) find that herding is present during market upswings. Bhaduri and Mahapatra (2013) report similar results, providing evidence that security return dispersion rises more slowly in up-markets than in down-markets, indicating herd-like behavior. Regarding China, Tan et al. (2008) report that herding occurs in both rising and falling markets, but that asymmetry is stronger during market upswings in the A-share market. No asymmetry is found in the B-share markets, which indicates that investor composition may explain the direction and strength of herding in China. Lao and Singh (2011), in turn, report that herding is greater during falling markets in China, which differs from the findings of Tan et al. On the other hand, Demirer and Kutan (2006) find no evidence of asymmetric herding in China. However, the authors find that return dispersions are lower during periods of high downside movement, suggesting that stocks move more similarly in declining markets. Overall, the evidence on asymmetric herding suggests that it exists, but the findings are mixed and vary across sample periods, markets, and investor groups.

3.5 Hypotheses

Building on prior literature, this section develops hypotheses regarding the existence and variation of market-wide herding in the Indian and Chinese stock markets during COVID-19. Prior research suggests that herding may be especially relevant in emerging equity markets, where information asymmetries and weaker firm-specific information may lead investors to rely more on broader market signals and to follow market consensus (Chiang & Zheng, 2010; Chang et al., 2000). At the same time, earlier findings for India and China are somewhat mixed, underscoring the importance of further empirical research on these markets. Furthermore, previous literature indicates that herding may not be constant over time, but may instead vary across different market conditions, such as extreme market events and rising or falling markets.

The null hypothesis of this thesis is based on the CSAD methodology of Chang et al. (2000). The authors suggest that rational asset pricing models imply a linear relationship between return dispersion and market returns. If investors follow the overall market behavior and ignore their own information, the linear relation between CSAD and market returns may not hold. Therefore, linearity is used as a benchmark for testing market-wide herding (Chang et al., 2000; Chiang & Zheng, 2010). Thus, the null hypothesis is the following:

H0: The relationship between CSAD and market returns is linear in the Indian and Chinese stock markets during the full sample period.

If the null hypothesis is rejected, it suggests that the CSAD-market return relationship is nonlinear, indicating a possibility for market-wide herding. The first hypothesis examines whether market-wide herding is present in the Indian and Chinese equity markets during the full sample period. Earlier studies have found herding in several emerging markets (Chang et al., 2000; Chiang & Zheng, 2010), and more specifically in the Indian and Chinese stock markets (Lao & Singh, 2011; Bhaduri & Mahapatra, 2013; Chiang et al., 2010; Tan et al., 2008). However, the evidence is not fully consistent, as Demirer and

Kutan (2006) find no evidence of herd behavior in the Chinese market. Given the characteristics of emerging markets and overall findings from prior studies, it is reasonable to expect that herding may be present in India and China during the full-sample period, 2.7.2018-30.6.2022. Based on that, the first hypothesis is as follows:

H1: Market-wide herding exists during the full sample period in the Indian and Chinese stock markets.

The second hypothesis considers whether herding varies across subperiods. The analysis is meant to test whether herding varies over time and in different market conditions in India and China. Earlier literature suggests that herding may depend on market conditions rather than remain constant over time. Christie and Huang (1995) argue that herding is more likely to emerge, especially under extreme market conditions, although they do not find supporting evidence. Chang et al. (2000) and Chiang and Zheng (2010), on the other hand, find that herding results vary across extreme market conditions and crisis periods. Regarding India and China specifically, Lao and Singh (2011) provide evidence of more pronounced herding during extreme market movements, while Demirer and Kutan (2006) find the opposite evidence in China. The evidence is therefore mixed, making it reasonable to expect that the intensity of herding may vary across subperiods, especially during the COVID-19 outbreak period, which was characterized by large market movements and increased uncertainty. Therefore, the second hypothesis is as follows:

H2: The extent of market-wide herding varies across subperiods in the Indian and Chinese stock markets.

The third hypothesis examines whether herding differs between up- and down-market days and, therefore, market states. Prior studies suggest that herding may be asymmetric across rising and falling markets, though the evidence remains inconclusive. Chang et al. (2000) find that herding during up- and down-markets differs across countries, while

Chiang and Zheng (2010) provide evidence that herding is present in both market directions and especially pronounced in Asia during rising markets. Furthermore, Lao and Singh's (2011) research suggests that herding is stronger during market upswings in India and during market downturns in China, whereas Chiang et al. (2010) and Tan et al. (2008) find that herding occurs in both rising and falling markets in China. Overall, the evidence is mixed, but it suggests that herding may be asymmetric between the two market directions in India and China. Based on the presented literature, the third hypothesis is as follows:

H3: The extent of market-wide herding differs during up- and down-market days in the Indian and Chinese stock markets during the full sample period

4 Data and descriptive statistics

The data comprises daily prices for constituents of the NSE Nifty 100 index in India and the Shanghai Shenzhen CSI 300 index in China. The Nifty 100 Index is a free-float capitalization-weighted index, meaning that index weights are proportional to constituents' tradable (free-float) market value. The index captures India's large-cap segment of publicly listed companies and, as of the time of writing, covers around 65% of the NSE (National Stock Exchange of India) free-float market capitalization, therefore providing a valid choice for a herding study on the Indian market. The Index consists of 100 companies and has a base value of 1000. The CSI 300 is a free-float, capitalization-weighted index composed of large, liquid A-shares (mainland China companies) listed on both the Shanghai and Shenzhen stock exchanges. The index consists of 300 companies and has a base value of 1000. The CSI 300 is commonly used as a proxy for China's onshore equity market.

In this thesis, both the Nifty 100 and the CSI 300 are chosen for a CSAD herding study because they provide broad target-market representation and have large, liquid constituents that reduce noisy dispersion driven by illiquidity, which improves the reliability of daily dispersion. Both indices include enough companies to yield valid results in a CSAD herding study, avoiding results dominated by a few individual returns. The CSI 300 also provides broader coverage of mainland China and, therefore, better coverage of the overall Chinese market than exchange-specific coverage.

Daily equity data for the indices were obtained from Refinitiv Datastream. For both indices, the dataset includes daily adjusted closing prices for index constituents, avoiding potential artificial returns from stock splits or corporate actions. The data sample period is from 02.07.2018 to 30.06.2022. All prices from both indices are reported in local currencies. For the Nifty 100, the currency is the Indian rupee (INR), and for the CSI 300, the currency is the Chinese renminbi (CNY). The original constituent-level data contained missing price data for some companies during the sample period. To construct a balanced panel and maintain a comparable sample of CSADs, firms without complete price histories were excluded from the data. Furthermore, trading days on which all constituent returns were zero were removed from the data because these days are non-trading days, which would artificially increase the number of zero-return days and downward-bias the CSAD measure. After these modifications, the final sample includes 94 companies for the Nifty 100 and 252 for the CSI 300. The number of daily observations for India is 989, and for China is 971. The observation amount varies between India and China due to different national holidays and trading days. Before calculating the daily CSADs of the stocks, the logarithmic daily returns are calculated with the following formula:

$$R_{i,t} = \ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right) \tag{12}$$

where $R_{i,t}$ is the daily logarithmic return of stock *i* on trading day t, $\ln$ denotes the natural logarithm, $P_{i,t}$ represents the adjusted closing price of a stock at time t, and $P_{i,t-1}$ refers to the adjusted closing price of a stock on the previous trading day t – 1. After calculating the daily returns for each index constituent, the daily market return is computed as the equally weighted average of those returns. The descriptive statistics for CSAD and equal-weighted market returns are presented in Table 1.

Table 1. Descriptive statistics for CSAD and equally weighted market returns for India and China

Market	**India**		**China**	
	CSAD	*Rm*	*CSAD*	*Rm*
Mean	0,0136	0,0005	0,0163	0,0006
Median	0,0128	0,0018	0,0156	0,0007
Standard Deviation	0,0045	0,0130	0,0041	0,0137
Kurtosis	11,6459	16,4396	0,7001	3,4139
Skewness	2,6136	-1,7003	0,8623	-0,5461
Minimum	0,0044	-0,1337	0,0084	-0,0862
Maximum	0,0512	0,0710	0,0332	0,0550
Observations	989	989	971	971

This table presents descriptive statistics for daily cross-sectional absolute deviations (CSAD) and daily equally weighted market returns for the NSE Nifty 100 and Shanghai Shenzhen CSI 300 indexes in India and China. The data period is 02.07.2018-30.06.2022

The descriptive statistics in Table 1 show that the mean daily returns of both markets are low, close to zero. Both India and China have roughly the same market return standard deviations (0,0130 and 0,0137). India's minimum daily market return (-0,1337) compared to the maximum (0,0710) indicates greater downside risk. China’s minimum (-0,0862) and maximum (0,0550) market returns also indicate a stronger downside risk, though less pronounced than in India. India has a higher kurtosis (16,4396), indicating a greater tendency for extreme returns and, therefore, higher tail risk, than China (3,4139). Both return distributions are negatively skewed, which suggests that large negative returns are more frequent than large positive returns. This is pronounced more strongly in India (-1,7003) than in China (-0,5461).

Regarding CSADs, Table 1 shows that China’s mean (0,0163) and median (0,0156) CSADs are higher than India’s (0,0136 and 0,0128), indicating that CSI 300 companies deviate more from the market return than Nifty 100 companies. Both markets have similar CSAD standard deviations (0,0041 and 0,0045). India has a high kurtosis (11,6459) compared to China (0,7001), suggesting a greater tendency for extreme dispersion events. India also has a higher CSAD skewness (2,6136) than China (0,8623). India’s CSADs range from 0,0044 to 0,0512, further suggesting more extreme dispersion than China’s 0,0084 to 0,0332.

As mentioned in the research hypothesis section, this thesis examines herding across three sub-periods to determine whether market-wide herding varies between extreme and normal market conditions. The pre-COVID period is between 2.7.2018 and 29.1.2020, outbreak period between 30.1.2020 and 30.6.2020, and post-outbreak period between 1.7.2020 and 30.6.2022. Figure 5 shows the index price data for Nifty 100 and CSI 300, divided into the three subperiods. The main period of interest is the outbreak period highlighted in red. The starting date was chosen based on the World Health Organization's declaration of COVID-19 as a global pandemic on 30.1.2020. The figure shows a clear market drawdown reaction in the Indian market and a more contained decline in the Chinese market during the outbreak period. The tightening of regulations in China's platform economy began in late 2020, and the "common prosperity" agenda, which started in mid-2021, might explain the weaker performance of the CSI 300 during the post-outbreak period. The start of the war in Ukraine in February 2022 might also explain some volatility in the late post-outbreak period for both indices. While both markets exhibit a sharp upward trend after the outbreak period, price movements across the three subperiods suggest substantial differences between the Indian and Chinese markets.

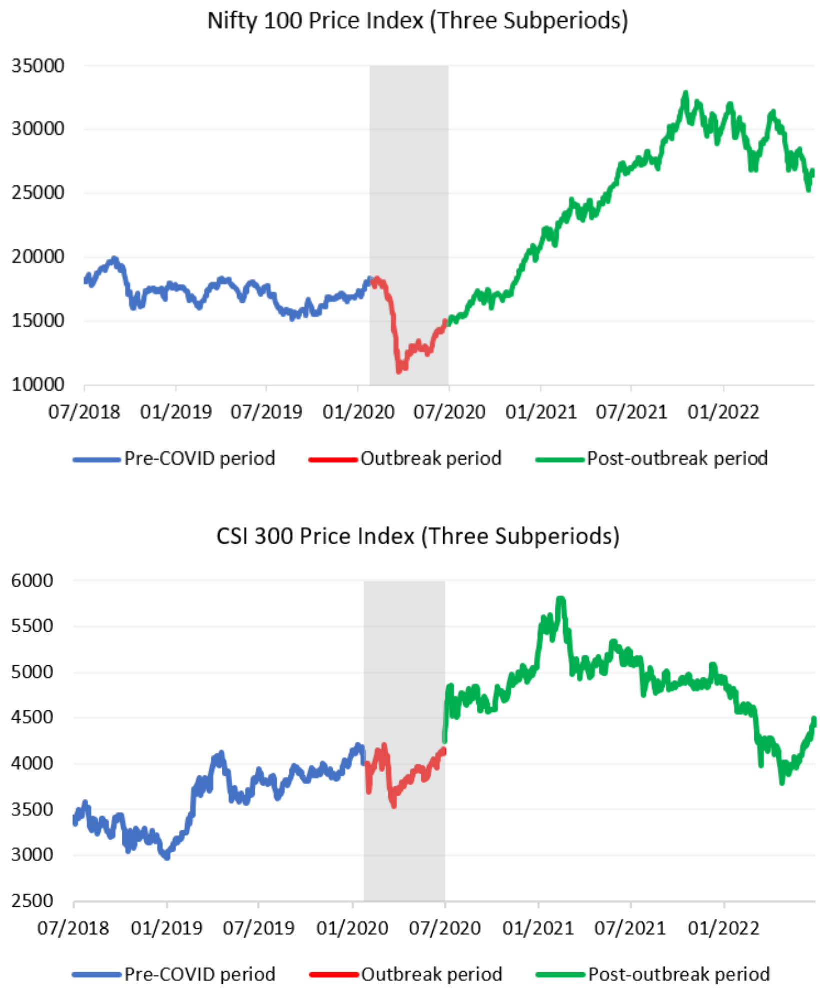

Figure 5. Nifty 100 and CSI 300 price index data divided into three subperiods

Figure 6 shows the relationship between CSAD and the magnitude of market returns in India and China. In both markets, the scatter shows a roughly V-shaped pattern, meaning that dispersion is low when market returns are close to zero and increases when the magnitude of market returns increases. The figure is consistent with the standard prediction that return dispersion tends to increase during larger market movements. China shows slightly higher return dispersion on average, whereas India shows more pronounced spikes in CSAD during large positive and negative market returns. However, these graphical observations are descriptive only, and the presence of herding cannot be concluded from them. The next chapter will present empirical methods intended to detect whether herding is present.

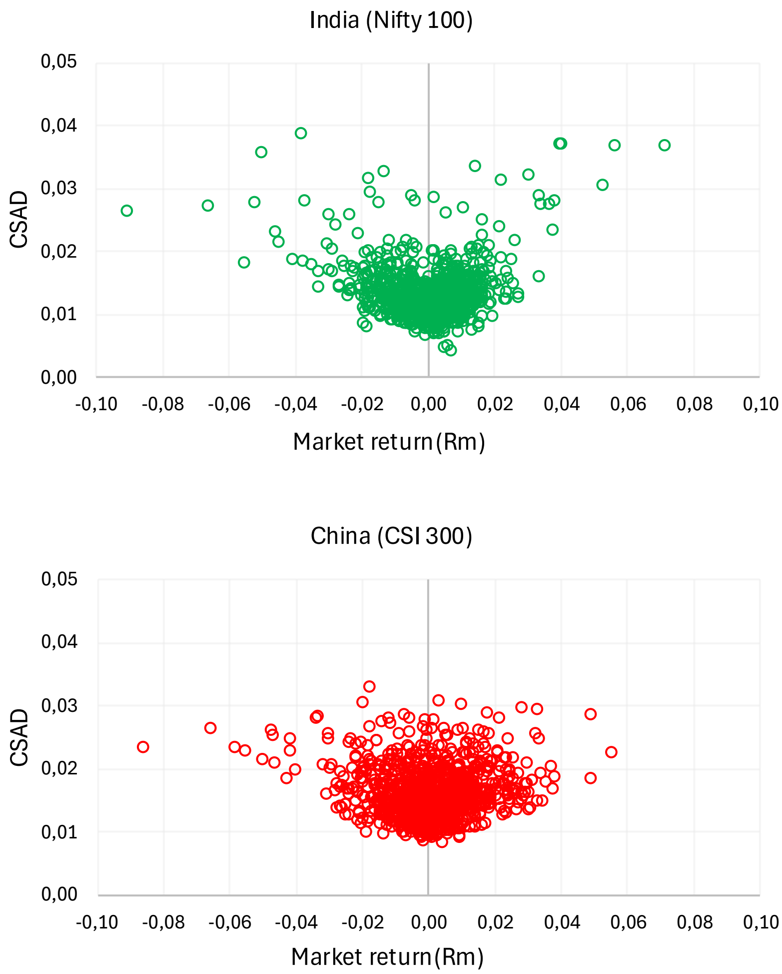

Figure 6. Nifty 100 and CSI 300 relationship between daily CSAD and equally weighted market return

5 Methodology

As noted in the literature review, this thesis uses cross-sectional data on stock returns to examine market-wide herding. Two main methods for that kind of herding measure have been proposed by Christie and Huang (1995) and Chang et al. (2000). Christie and Huang argued that investors' investment decisions vary with market conditions. During normal market periods, rational asset-pricing models predict that cross-sectional asset dispersion will increase as absolute market returns rise, because investors make decisions based on their own, diverse information. During periods of extreme volatility, however, investors may ignore their private information and align their decisions with the broader market. As a result, individual stock returns tend to cluster around the market return. The authors therefore propose the cross-sectional standard deviation method to measure the dispersion of returns. Because CSSD is based on squared return deviations, it is highly sensitive to outliers. Chang et al., therefore, propose the original cross-sectional absolute deviation method, which forms the basis of the methodology used in this thesis. To begin, the CSAD is defined with the following formula:

$$CSAD_t = \frac{1}{N}\sum_{i=1}^{N}\left|R_{i,t} - R_{m,t}\right| \tag{13}$$

Where $CSAD_t$ is the cross-sectional absolute deviation of stock returns at time t, N is the overall number of stocks in the portfolio, $R_{i,t}$ is the return of stock *i* on day t, and $R_{m,t}$ is the equal-weighted market portfolio return. Based on the original CSAD regression equation by Chang et al. (2000), Chiang and Zheng (2010) add an additional $R_{m,t}$ term to the equation, which accounts for asymmetric investor behavior across different market states. After calculating the daily CSADs, the slightly modified regression equation by Chiang and Zheng (2010) is employed to examine whether herding is detected:

$$CSAD_t = \gamma_0 + \gamma_1 R_{m,t} + \gamma_2\left|R_{m,t}\right| + \gamma_3 R_{m,t}^2 + \varepsilon_t \tag{14}$$

Where, in addition to equation (13), γ_0 is the constant term, γ_1 is the coefficient on the signed market return $R_{m,t}$, γ_2 is the coefficient on the absolute market return $|R_{m,t}|$, γ_3 is the coefficient on the squared market return $R^2_{m,t}$, and ε_t is the error term. γ_3 is the main coefficient of interest in the equation. As mentioned earlier, rational asset-pricing models predict a linear relationship between the dispersion of individual stock returns and the market portfolio return, meaning that as the absolute market return increases, cross-sectional return dispersion should also rise. Therefore, a negative and significant γ_3 coefficient suggests that return dispersion increases at a decreasing rate during large market moves, indicating a non-linear relationship, and that stock return dispersion clusters around the market return during high market volatility. A negative and significant γ_3 would thus be evidence of herding.

As noted in the literature review, many studies have found that herding may vary across different market states (Chiang & Zheng, 2010; Tan et al., 2008; Hwang & Salmon, 2004). To test whether investors behave differently during up- and down-market days, the data is divided into two groups using a dummy variable, as in Chiang and Zheng (2010). The equation is the following:

$$CSAD_t = \gamma_0 + \gamma_1(1-D)R_{m,t} + \gamma_2 DR_{m,t} + \gamma_3(1-D)R^2_{m,t} + \gamma_4 DR^2_{m,t} + \varepsilon_t \quad (15)$$

Where, in addition to equation (14), D is a dummy variable. D is defined as a market-state indicator where it equals one if $R_{m,t}$ is negative. Likewise, D equals to zero if $R_{m,t}$ is positive. Therefore, a negative and statistically significant γ_3 indicates herding behavior during up-market days, and similarly, a negative and statistically significant γ_4 implies herding on down-market days.

As this study relies on time-series data, autocorrelation may be present, which can affect coefficient estimation and lead to biased results. To mitigate this issue, all regressions are estimated using an AR(1) specification, following Lao & Singh (2011). This helps to reduce potential first-order autoregressive serial correlation and improve the reliability of the results. For robustness and comparison, the regressions are also estimated without the AR(1) specification and are reported in the appendices.

6 Empirical results

This chapter presents the empirical findings of the thesis using three complementary regression models. The first tests whether market-wide herding exists over the entire sample period. The second assesses whether herding varies across the pre-COVID, outbreak, and post-outbreak subperiods. The third examines whether herding varies between up-market and down-market conditions. Overall, these tests enable a broader evaluation of whether herding is present, how it changes over time and market conditions, and whether it depends on market direction.

6.1 Herding during the full sample period

Several papers have found herding in emerging and Asian equity markets (Chiang & Zheng, 2010; Chang et al., 2000; Tan et al., 2008; Lao & Singh, 2011). Therefore, the first empirical test, based on equation (14), examines whether herding exists in the Indian and Chinese markets over the full sample period. This provides a baseline test of herding before testing whether herding is affected by subperiods or market states. The first regression tests for the presence of herding during the full sample period in both the Indian and Chinese stock markets.

Table 2 reports the results of the CSAD regressions for India and China over the full sample period. As mentioned before, in the CSAD framework, herding is indicated when the coefficient γ_3 on squared market return is negative and statistically significant, indicating a non-linear relationship between stock return dispersion and market return, contrary to the predictions of rational asset pricing models (Chiang & Zheng, 2010).

Table 2. Market-wide herding analysis over the full sample period

Estimated coefficients for market-wide herding over the full sample period						
	γ_0	γ_1	γ_2	γ_3	**AR(1)**	Adj. R^2
Market						
India	0,0123***	-0,0014	0,1437***	-0,0021	0,6628***	0,6024
	(36,5300)	(-0,3324)	(13,1542)	(-0,0180)	(42,0729)	
China	0,0152***	-0,0100*	0,1237***	-0,5096**	0,6778***	0,5152
	(44,2213)	(-1,7352)	(9,0120)	(-1,9982)	(29,7078)	

This table reports the estimated coefficients of equation (14) for the full sample period 02.07.2018-30.6.2022, estimated with an AR(1) specification to account for autocorrelation. A negative and statistically significant coefficient on γ_3 indicates herding. The T-statistics are reported in parentheses. Statistical significance is reported as follows:
* Statistical significance at the 10% level
**Statistical significance at the 5% level
***Statistical significance at the 1% level

Table 2 shows that the full-sample regression results exhibit a negative but statistically insignificant coefficient for India (-0,0021), indicating no evidence of herding over the full sample period. Thus, the Indian market appears more consistent with rational asset pricing models than with herding during the sample period. However, in China, the γ_3 coefficient is negative and statistically significant (-0,5096**), which provides evidence of herding. This indicates that return dispersion in China increases at a decreasing rate during large market movements, consistent with herding behavior. Therefore, the null hypothesis is rejected for the Chinese market, and Hypothesis 1 is supported for China, but not for India.

The differing results between India and China may be partly explained by the investor structures of these markets. Lao and Singh (2011) suggest in their study, which finds herding in both India and China, that the stronger herding results for China may be due to an increased number of individual investors, who are typically considered inexperienced. They also suggest that the smaller amount of herding in India in their study may be due to the stronger influence of institutions, which are typically associated with greater skills and better information. Therefore, the increased number of individual investors in China and the stronger impact of institutional investors in India may also explain the different results in this thesis. Furthermore, as the B-share market in China is associated with foreign institutional investors (Tan et al., 2008), comparing Chinese B-shares with the Indian equity market for the same sample period could provide evidence on the impact of investor composition on herding in these markets.

As a robustness test, the analysis is estimated without an AR(1) term, which controls for autocorrelation, and the results differ significantly, as reported in Appendix 1. The coefficients without the AR(1) specification do not provide evidence of herding in either India or China, which suggests that the full sample findings, especially for China, are sensitive to autocorrelation.

The results indicating no herding in India and herding in China are partially consistent with earlier studies covering the same markets. Lao and Singh (2011) also find evidence of herding in the Chinese markets using the CSAD approach for the period 1999-2009. However, they also document herding in India, which differs from the findings of this thesis. One possible explanation for this difference is the different sample period. Regarding China, the results are consistent with Tan et al. (2008), who find herding in Shanghai and Shenzhen A-shares during 1997-2003, but they differ from those of Demirer and Kutan (2006), who find no evidence of herding in China during 1999-2002. Overall, the full-sample regression results are consistent with the literature that finds no herding in India and discovers evidence of herding in China. However, herding might be time-varying, depending on market conditions or market state, and therefore further analysis of herding across subperiods and market states is necessary to provide a more robust result.

6.2 Herding during different subperiods

Prior literature suggests that herding might be pronounced during periods of market stress (Chiang & Zheng, 2010). Furthermore, some papers suggest that herding is condition-dependent and may vary across different market environments (Hwang &

Salmon, 2004; Tan et al., 2008). Therefore, the second regression tests whether market-wide herding varies across different subperiods.

Table 3 reports the results of the CSAD regression for India and China over the three subperiods: pre-COVID, outbreak, and post-outbreak. The subperiod analysis examines whether herding varies across market conditions and time periods in the selected markets. The pre-COVID period represents a relatively normal market, while the outbreak period reflects highly volatile market conditions during the crisis. Assuming herding exists and that it is pronounced during extreme market movements, it should be evident at least during the outbreak period. Lastly, the post-outbreak period represents a period of normalization after the start of the crisis. The same equation (14) is used here as in Table 2, but now it is tested for the different subperiods. The main coefficient of interest, referring to the squared market return, in Table 3, is γ_3, which must be negative and statistically significant for herding to exist during the specific subperiods.

Table 3. Market-wide herding analysis over the three subperiods

Estimated coefficients for market-wide herding over the three subperiods						
Panel A: Pre-COVID period						
	γ_0	γ_1	γ_2	γ_3	AR(1)	Adj. R^2
Market						
India	0,0119***	-0,0003	0,0883***	3,8729***	0,5109***	0,4714
	(38,2518)	(-0,0251)	(3,2613)	(6,5629)	(14,7566)	
China	0,0127***	-0,0165**	0,1308***	-0,1399	0,4957***	0,4286
	(37,0728)	(-2,2064)	(4,7589)	(-0,2415)	(11,3243)	
Panel B: Outbreak period						
	γ_0	γ_1	γ_2	γ_3	AR(1)	Adj. R^2
India	0,0176***	-0,0051	0,1433***	-0,2029	0,7475***	0,7030
	(8,3999)	(-0,2750)	(3,6269)	(-0,5298)	(8,1474)	
China	0,0143***	-0,0228	0,2063***	-1,7720**	0,4879***	0,3789
	(17,7653)	(-1,0384)	(4,8325)	(-2,2102)	(5,5389)	
Panel C: Post-outbreak period						
	γ_0	γ_1	γ_2	γ_3	AR(1)	Adj. R^2
India	0,0119***	-0,0110	0,0864**	1,4415	0,4030***	0,2939
	(39,4231)	(-1,4189)	(2,5430)	(1,6079)	(9,9714)	
China	0,0172***	-0,0028	0,1072***	-0,4546	0,6145***	0,4282
	(41,0325)	(-0,2946)	(3,8734)	(-0,6266)	(16,8337)	

This table reports the estimated coefficients of equation (14) for the three subperiods estimated with an AR(1) specification to account for autocorrelation. A negative and statistically significant coefficient on γ_3 indicates herding. The T-statistics are reported in parentheses. The subperiods and statistical significance are as follows:
Pre-COVID period: 02.07.2018-29.1.2020
Outbreak period: 30.1.2020-30.6.2020
Post-outbreak period: 01.07.2020-30.6.2022
* Statistical significance at the 10% level
**Statistical significance at the 5% level
***Statistical significance at the 1% level

Based on Table 3, neither India nor China shows herding behavior during the pre-COVID period. India exhibits strong evidence against herding, with a positive and statistically significant coefficient (3,8729***), which suggests that return dispersion increases during periods of large market movements. In China, the γ_3 coefficient is negative but statistically insignificant. During the outbreak period, both markets have a negative γ_3 coefficient, suggesting that return dispersion is directionally consistent with herding. However, for India, the evidence is too weak to conclude that herding is present. For China, the coefficient is statistically significant (-1,7720**), which indicates that herding occurred during the outbreak period. This finding is noteworthy, as it supports the view that crisis periods are often associated with increased uncertainty, during which herding may be more likely to emerge. In the post-outbreak period, neither market shows statistically significant evidence of herding, although the coefficient for China remains negative. Overall, Table 3 indicates no evidence of herding for India across the three subperiods, whereas China exhibits herding during the outbreak period. Therefore, Hypothesis 2 is supported for China, but not for India.

Since Chinese A-shares are typically associated with less-informed individual investors (Tan et al., 2008), and the outbreak period was a period of market decline, the herding result for China during the outbreak period may reflect individual investors following the overall market and seeking to avoid potential losses during a crisis period. Therefore, the different results between India and China could potentially be explained by differences in investor composition.

The results differ when the regression is estimated without the AR(1) specification, as reported in Appendix 2. Without the AR(1) term, neither market shows evidence of herding in any of the three subperiods, although both markets show negative but statistically insignificant coefficients during the outbreak period. This suggests that the subperiod findings, particularly for China, are sensitive to autocorrelation.

The results on how herding varies with time and market conditions are partly consistent with some of the earlier literature on herding. Regarding the Chinese market, Tan et al. (2008) find that herding is higher during high volume and volatility. This thesis suggests somewhat similar results, as the outbreak period was highly volatile and herding was present in China. On the other hand, the results differ from Demirer and Kutan (2006), who find no herding in China during periods of high volatility. Regarding India, Lao and Singh (2011) and Bhaduri and Mahapatra (2013) find that herding is more pronounced during extreme market conditions. This thesis finds different results, as herding is not found in India during the outbreak period.

6.3 Herding during up- and down-markets

The standard CSAD regression examines whether herding exists on average in the market, but it does not differentiate between rising and falling markets. Investors may behave differently across market states, as prior literature on herding in up- and down-markets indicates (Chiang & Zheng, 2010; Lao & Singh, 2011; Tan et al., 2008). Therefore, the third regression tests whether herding varies across rising and falling markets.

Table 4 reports the results of the CSAD regression for India and China during up- and down-market days over the full sample period. The regression is used to examine whether there is herding asymmetry across different market states. Using equation (15) by Chiang and Zheng (2010), the test focuses on the non-linear coefficients γ_3 and γ_4 which capture herding during up- and down-market days. A negative and statistically significant γ_3 coefficient on squared market return implies herding during up-market days, and similarly, a negative and statistically significant γ_4 indicates herding during down-market days.

Table 4. Market-wide herding analysis during up- and down-market days

Estimated coefficients for market-wide herding during up- and down-market days over the full sample period							
	γ_0	**γ_1**	**γ_2**	**γ_3**	**γ_4**	**AR(1)**	Adj. R^2
Market							
India	0,0124***	-0,1018***	-0,1449***	1,2746***	-0,0361	0,6589***	0,6044
	(36,6935)	(5,2938)	(-10,9448)	(3,2100)	(-0,3240)	(39,9322)	
China	0,0152***	0,1059***	-0,1347***	-0,2130	-0,5604**	0,6780***	0,5147
	(44,0619)	(4,7412)	(-8,2799)	(-0,3085)	(-2,1795)	(29,7046)	

This table reports the estimated coefficients of equation (15) for the full sample period 02.07.2018-30.6.2022 estimated with an AR(1) specification to account for autocorrelation. A negative and statistically significant coefficient on γ_3 indicates herding during up-market days while a negative and statistically significant coefficient on γ_4 indicates herding during down-market days. The T-statistics are reported in parentheses. Statistical significance is reported as follows:
* Statistical significance at the 10% level
**Statistical significance at the 5% level
***Statistical significance at the 1% level

Table 4 shows that India does not exhibit herding behavior in either up- or down-market days over the full sample period. On up-market days, India's γ_3 coefficient is positive and statistically significant (1,2746***), indicating that return dispersion increases at an increasing rate, which is inconsistent with herding. In China, both the up- and down-market coefficients are negative, but only the down-market γ_4 coefficient is statistically significant. This provides evidence of herding during down-market days over the full sample period and suggests that herding is asymmetric with respect to market direction. Therefore, Hypothesis 3 is supported for China, but not for India. The different results for India and China during falling markets could be partly explained by investor composition, as individual investors may be more inclined to avoid losses in declining markets (Lao & Singh, 2011), and the Chinese A-shares are typically associated with individual investors, as mentioned before.

The results differ again when the regression is done without the AR(1) specification, as reported in Appendix 3. Without the AR(1) term, neither market shows statistically significant evidence of herding during up- or down-market days. While China has negative coefficients that are directionally consistent with herding, the evidence is too weak to conclude that herding is present. Therefore, the results from Appendix 3 suggest that the findings, especially for China, are sensitive to autocorrelation.

Furthermore, as a robustness check, the simplified up- and down-market specification introduced by Chang et al. (2000) and subsequently applied by Tan et al. (2008) for the Chinese market is estimated. Unlike the main specification, this approach separates the analysis into two regressions, one for up-days and one for down-days. The results are broadly consistent for India and support the findings in Table 4. However, the results differ materially for China. In the main model, China's nonlinear coefficient on down-market days is negative and statistically significant, whereas in the simplified specification, it is positive and insignificant. This suggests that the evidence of asymmetric herding in China is sensitive to model specification.

The results on herding during up- and down-markets are partly consistent with earlier literature. For India, the findings differ from Lao and Singh (2011), who find more pronounced herding during up-market trends. As for China, however, the authors find that herding is more pronounced during falling markets, which is consistent with the results of this thesis. Similarly, Tan et al. (2008) find evidence of herding in both rising and falling markets in China, which is partially consistent with this thesis's results, which indicate herding on down-market days. On the other hand, Demirer and Kutan (2006) find no evidence of asymmetric herding in China, a finding that differs from this thesis. However, the authors find that return dispersions are lower during periods of high downside movement, suggesting that stocks move more similarly in declining markets. Overall, the results of this thesis support the view that herding in China may be asymmetric and dependent on market conditions. For India, the findings are more consistent with the predictions of rational asset pricing models.

6.4 Limitations

There are several limitations related to the CSAD methodology and, therefore, also how the results of this thesis should be interpreted. As mentioned before, the CSAD itself is not a direct measure of herding, but herding is inferred from the relationship between return dispersion and market returns (Chang et al., 2000). This means that the method does not directly observe investors' actual trading motives or whether they ignored their own information and followed the market consensus (Spyrou, 2013). Since it is difficult to directly examine whether investors ignore their own information without detailed data on investors' information before a trade, many empirical studies instead examine herding at the aggregate market level. Furthermore, although a negative and statistically significant nonlinear coefficient is taken as evidence of herding using the CSAD methodology, it is not definitive proof of herding. Chang et al. (2000) note that the nonlinear relationship between the CSAD and market returns could also arise from a nonlinear market model.

Market-wide herding measures, such as the CSAD method, are typically based on return dispersion. Therefore, they capture only one specific form of herding, which is herding toward the market consensus (Spyrou, 2013). While this thesis does not find herding at the market level, this does not mean that herding does not exist within certain sectors, industries, or investor groups. Therefore, the absence of market-wide herding does not necessarily imply that herding is completely absent in the Indian market, but it indicates that herding is not present at the aggregate market level. Furthermore, for China, this means that herding can also exist across different sectors and investor groups during the sample period.

The methodology used also does not distinguish between the causes of herd behavior, such as whether herding is spurious or intentional, as described by Bikhchandani and Sharma (2001). In spurious herding, investors make similar decisions because they have access to the same information, such as macroeconomic news. Intentional herding refers to a situation in which investors intentionally copy one another, for instance, to maintain their reputation in the job market. Therefore, similar movement in stock returns may indicate imitation and herding, but it may also result from investors reacting to changes in fundamental information. Hwang and Salmon (2004) emphasize the importance of the ability to empirically distinguish between spurious and intentional herding, as the former reflects an efficient reaction to new information, whereas the latter may lead to inefficiency and mispricing. Furthermore, Spyrou (2013) notes that empirical herding methodologies typically examine clustering in trades and decisions, but this clustering may arise for various unexplained reasons. This limitation is also relevant to the CSAD approach, as it interprets herding indirectly via the relationship between return dispersion and market returns. Therefore, it does not explain the reasons for potential herding. The CSAD method also does not explain whether herding is rational or irrational.

A further limitation for the results of this thesis is that domestic return dispersion may be influenced by events in major foreign markets. Chiang and Zheng (2010) argue that herding models restricted to a single domestic market may exclude relevant external influences and may therefore lead to biased estimates. This is especially relevant in integrated financial markets, where shocks may not remain confined to a single region or country. According to the authors, the role of the US market, in particular, should not be overlooked when analyzing herding in a single market.

Spyrou (2013) further notes that herding may vary over time, and many empirical studies do not examine whether the same factors drive herding behavior in different periods or whether the same market participants are responsible for it. This suggests that a full-sample analysis may conceal differences across time and market conditions. For this reason, this thesis examined herding across three COVID-19-related subperiods and across different market states. While this approach helps account for possible time variation in herding, it does not explain whether the underlying causes of herding or the market participants who herd remain the same over time.

7 Conclusions

This thesis examined market-wide herding in the Indian and Chinese equity markets during the COVID-19 period using the CSAD methodology introduced by Chang et al. (2000) and, more specifically, the modified specification proposed by Chiang and Zheng (2010), estimated with an AR(1) specification following Lao and Singh (2011). The empirical results show clear differences between the Indian and Chinese equity markets. The full-sample regression results do not indicate statistically significant herding in India. For China, the nonlinear coefficient is negative and statistically significant, indicating herding over the full-sample period. The subperiod regression analysis yields similar results for India, with no statistically significant herding found in any subperiod. In contrast, China shows a negative and statistically significant coefficient during the outbreak period, suggesting that herding was present during a period of high COVID-19-related market uncertainty.

The up- and down-market regression analysis also shows differences between India and China. For India, no statistically significant herding is detected during rising or falling market days. Furthermore, the coefficient for up-market days is positive and statistically significant, indicating evidence against herding behavior. For China, the nonlinear coefficient for down-market days is negative and statistically significant, which suggests that market-wide herding occurred during falling market conditions. However, this finding is not observed when the up- and down-market regression analyses are estimated separately following Chang et al. (2000), indicating that the evidence of asymmetric herding in China is sensitive to model specification. In addition, all three main regressions are also estimated without the AR(1) term, following Chiang and Zheng (2010), and no statistically significant herding is found in either market in any of the regressions without the AR(1) term. This suggests that the findings, especially for China, are sensitive to whether autocorrelation is controlled for with the AR(1) specification.

This thesis's findings for India indicate that individual stock returns did not cluster around the market return in a way consistent with market-wide herding during large market movements in the COVID-19 sample period. Instead, the results suggest that the Indian equity market behaved more consistently with the predictions of rational asset pricing models. The findings differ for China, where herding is found in the full-sample regression, during the outbreak period, and during down-market days. Therefore, the findings from the Chinese equity market support the view that herding may become more pronounced under extreme market conditions and may vary with market state.

The results should be interpreted with several limitations. First, the CSAD methodology does not directly observe investor behavior or trading motives. Rather, herding is examined on the market level by investigating the relationship between return dispersion and market returns (Chang et al., 2000). Therefore, for India, the absence of statistically significant herding should be interpreted as evidence against market-wide herding rather than as proof that herding is entirely absent during the sample period. Second, market-wide herding measures capture herding towards the market consensus (Spyrou, 2013), meaning that herding may still exist within specific sectors, industries, or investor groups. This is also relevant for China, where the results indicate market-wide herding, but do not identify which investors or market segments may drive the behavior. Third, the CSAD approach does not explain the reasons behind possible market-wide herding. Therefore, the methodology does not explain whether herding is caused by intentional imitation or by spurious herding driven by rational investor reactions to common information. Finally, the analysis focuses on returns in the Indian and Chinese domestic markets and does not consider foreign market influences, which may be relevant in an increasingly integrated financial market (Chiang & Zheng, 2010).

The results have several practical implications. For investors and portfolio managers, the findings suggest different implications for the Indian and Chinese equity markets. In India, portfolio diversification benefits do not appear to be substantially reduced by market-wide herding. In China, however, the results suggest that individual stock returns clustered more around the market return during large market movements. This may reduce diversification benefits during periods of market stress, when diversification is usually important. Furthermore, the findings suggest that herding-related mispricing may be more relevant in the Chinese equity market than in the Indian equity market. Regarding practical implications for regulators, previous literature on emerging markets suggests that weaker regulation, lower information disclosure standards, limited firm-specific information, and less-informed investors may contribute to herding (Chang et al., 2000; Demirer & Kutan, 2006). The evidence of herding in China, therefore, suggests that regulators should continue to support market transparency, reliable disclosure of firm-specific information, and investor protection. The same applies to India, as the results do not rule out herding within specific sectors or investor groups.

Future research could extend this thesis in several ways. First, since the results differ between India and China, future studies could examine whether similar differences exist in other emerging markets, including smaller, less-developed ones where herding may be more likely to occur. Second, as this thesis focuses on market-wide herding, future studies could investigate whether herding is present at the sector or industry level. This would help to determine whether herding exists within specific parts of the market. Third, future research could investigate whether herding differs across investor groups, such as retail, institutional, or foreign investors, since different investor types may react differently under market uncertainty. This would be particularly relevant in China, where A-shares, which were also examined in this thesis, have historically been associated more with domestic individual investors, while B-shares have been associated more with foreign institutional investors (Tan et al., 2008). Further research could also investigate whether differences in investor composition explain the differing herding results between India and China during the sample period. Fourth, research on Indian and Chinese equity markets could examine herding using alternative methodologies and model specifications to assess whether findings on market-wide herding vary by choice of methodology. This is particularly relevant because evidence of herding in China in this thesis relies on the inclusion of the AR(1) specification.

Furthermore, since several limitations remain in the existing herding measures, future research should continue to develop better empirical methodologies for detecting herding. As Spyrou (2013) notes, existing empirical evidence remains inconclusive, partly due to methodological limitations and because different methods test for different forms of herding. It is therefore important to clearly define what form of herding is being tested and to develop appropriate measures for examining that form of herding. As herding may be short-lived (Christie & Huang, 1995; Lao & Singh, 2011), future research could also examine whether herding occurs during shorter periods of market stress, including recent trade-policy or geopolitical shocks, such as the 2025 “Liberation Day” tariff-related market downturn.

References

Akbar, Y. H., & Samii, M. (2005). Emerging markets and international business: A research agenda. Thunderbird International Business Review, 47(4), 389. https://doi.org/10.1002/tie.20059

Baddeley, M. C., Curtis, A., & Wood, R. (2004). An introduction to prior information derived from probabilistic judgements: elicitation of knowledge, cognitive bias and herding. Geological Society, London, Special Publications, 239(1), 15-27. https://doi.org/10.1144/gsl.sp.2004.239.01.02

Baker, H. K., & Ricciardi, V. (2014). Investor behavior: The psychology of financial planning and investing. John Wiley & Sons.

Banerjee, A. V. (1992). A simple model of herd behavior. The Quarterly Journal of Economics, 107(3), 797-817. https://doi.org/10.2307/2118364

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. Journal of Financial Economics, 49(3), 307-343. https://doi.org/10.1016/S0304-405X(98)00027-0

Bhaduri, S. N., & Mahapatra, S. D. (2013). Applying an alternative test of herding behavior: A case study of the Indian stock market. Journal of Asian Economics, 25, 43-52. https://doi.org/10.1016/j.asieco.2013.02.001

Bikhchandani, S., Hirshleifer, D. and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy, Vol. 100 (5), 992-1026. https://doi.org/10.1086/261849

Bikhchandani, S., & Sharma, S. (2001). Herd behavior in financial markets. IMF Staff Papers, 47(3), 279-310. https://doi.org/10.2307/3867650

Black, F. (1986). Noise. The Journal of Finance, 41(3), 528-543. https://doi.org/10.1111/j.1540-6261.1986.tb04513.x

Bodie, Z., Kane, A., & Marcus, A. J. (2023). Investments (Thirteenth edition). McGraw-Hill Education.

Caparrelli, F., D'Arcangelis, A. M., & Cassuto, A. (2004). Herding in the Italian stock market: a case of behavioral finance. The Journal of Behavioral Finance, 5(4), 222-230. https://doi.org/10.1207/s15427579jpfm0504_5

Chang, E. C., Cheng, J. W., & Khorana, A. (2000). An examination of herd behavior in equity markets: An international perspective. Journal of Banking & Finance, 24(10), 1651-1679. https://doi.org/10.1016/S0378-4266(99)00096-5

Chiang, T. C., Li, J., & Tan, L. (2010). Empirical investigation of herding behavior in Chinese stock markets: Evidence from quantile regression analysis. Global Finance Journal, 21(1), 111-124. https://doi.org/10.1016/j.gfj.2010.03.005

Chiang, T. C., & Zheng, D. (2010). An empirical analysis of herd behavior in global stock markets. Journal of Banking & Finance, 34(8), 1911-1921. https://doi.org/10.1016/j.jbankfin.2009.12.014

Choi, N., & Sias, R. W. (2009). Institutional industry herding. Journal of Financial Economics, 94(3), 469-491. https://doi.org/10.1016/j.jfineco.2008.12.009

Choijil, E., Méndez, C. E., Wong, W. K., Vieito, J. P., & Batmunkh, M. U. (2022). Thirty years of herd behavior in financial markets: A bibliometric analysis. Research in International Business and Finance, 59, 101506. https://doi.org/10.1016/j.ribaf.2021.101506

Christie, W. G., & Huang, R. D. (1995). Following the pied piper: do individual returns herd around the market?. Financial Analysts Journal, 51(4), 31-37. https://doi.org/10.2469/faj.v51.n4.1918

De Bondt, W. F., & Thaler, R. (1985). Does the stock market overreact?. The Journal of Finance, 40(3), 793-805. https://doi.org/10.1111/j.1540-6261.1985.tb05004.x

Demirer, R., & Kutan, A. M. (2006). Does herding behavior exist in Chinese stock markets?. Journal of International Financial Markets, Institutions and Money, 16(2), 123-142. https://doi.org/10.1016/j.intfin.2005.01.002

Demirer, R., Kutan, A. M., & Chen, C.-D. (2010). Do investors herd in emerging stock markets? Evidence from the Taiwanese market. Journal of Economic Behavior & Organization, 76(2), 283–295. https://doi.org/10.1016/j.jebo.2010.06.013

Devenow, A., & Welch, I. (1996). Rational herding in financial economics. European Economic Review, 40(3–5), 603–615. https://doi.org/10.1016/0014-2921(95)00073-9

Economou, F., Kostakis, A., & Philippas, N. (2011). Cross-country effects in herding behaviour: Evidence from four south European markets. Journal of International Financial Markets, Institutions and Money, 21(3), 443-460. https://doi.org/10.1016/j.intfin.2011.01.005

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. Journal of Financial Economics, 116(1), 1–22. http://dx.doi.org/10.1016/j.jfineco.2014.10.010

Fama, E. F., & French, K. R. (2018). Choosing factors. Journal of Financial Economics, 128(2), 234–252. https://doi.org/10.1016/j.jfineco.2018.02.012

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. Journal of Finance, 25(2), 28-30. https://doi.org/10.2307/2325486

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. the Journal of Finance, 47(2), 427-465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

Froot, K.A., Scharfstein, D.S. & Stein, J.C. (1992). Herd on the street: informational inefficiencies in a market with short-term speculation. The Journal of Finance, Vol. 47 (4), 1461-1484. https://doi.org/10.2307/2328947

Graham, J.R. (1999). Herding among investment newsletters: theory and evidence. The Journal of Finance, Vol. 54 (1), 237-268. https://doi-org.proxy.uwasa.fi/10.1111/0022-1082.00103

Grinblatt, M., Titman, S., & Wermers, R. (1995). Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. American Economic Review, 85(5), 1088-1105. Retrieved 25-9-2025 from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5686

Hirshleifer, D. (2015). Behavioral finance. Annual Review of Financial Economics, 7, 133–159. https://doi.org/10.1146/annurev-financial-092214-043752

Hirshleifer, D. (2001). Investor psychology and asset pricing. The Journal of Finance, 56(4), 1533-1597. https://doi.org/10.1111/0022-1082.00379

Hong, H., Kubik, J. D., & Solomon, A. (2000). Security analysts' career concerns and herding of earnings forecasts. The Rand Journal of Economics, 121-144. https://doi.org/10.2307/2601032

Hwang, S., & Salmon, M. (2004). Market stress and herding. Journal of Empirical Finance, 11(4), 585-616. https://doi.org/10.1016/j.jempfin.2004.04.003

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. Journal of Finance, 48(1), 65–91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x

Kahneman, D., & Tversky, A. Prospect Theory: An Analysis of Decision under Risk. Econometrica, 47(1), 263-292. https://doi.org/10.2307/1914185

Kendall, M. G., & Hill, A. B. (1953). The analysis of economic time-series-part i: Prices. Journal of the Royal Statistical Society. Series A (General), 116(1), 11-34. https://doi.org/10.2307/2980947

Keynes, J.M. (1936). The General Theory of Employment, Interest and Money. Macmillan Publications, London.

Khanna, T., & Palepu, K. (2000). Is group affiliation profitable in emerging markets? An analysis of diversified Indian business groups. The Journal of Finance, 55(2), 867-891. https://doi.org/10.1111/0022-1082.00229

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. Journal of Finance, 49(5), 1541–1578. https://doi.org/10.2307/2329262

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1992). The impact of institutional trading on stock prices. Journal of Financial Economics, 32(1), 23-43. https://doi.org/10.1016/0304-405X(92)90023-Q

Lao, P., & Singh, H. (2011). Herding behaviour in the Chinese and Indian stock markets. Journal of Asian Economics, 22(6), 495-506. https://doi.org/10.1016/j.asieco.2011.08.001

Li, G. (2008). China's stock market: Inefficiencies and institutional implications. China & World Economy, 16(6), 81-96. https://doi.org/10.1111/j.1749-124X.2008.00139.x

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Review of Economics and Statistics, 47(1), 13–37. https://doi.org/10.2307/1924119

Liu, T., Zheng, D., Zheng, S., & Lu, Y. (2023). Herding in Chinese stock markets: Evidence from the dual-investor-group. Pacific-Basin Finance Journal, 79, 101992. https://doi.org/10.1016/j.pacfin.2023.101992

Neumann, J., & Morgenstern, O. (1944). Theory of Games and Economic Behavior. Princeton University Press.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. Journal of Economic Perspectives, 17(1), 59-82. Retrieved 9-10-2025 from https://www-jstor-org.proxy.uwasa.fi/stable/3216840?seq=2

Markowitz, H. (1952). Portfolio selection. The Journal of Finance, 7(1), 77–91. https://doi.org/10.2307/2975974

Mobarek, A., Mollah, S., & Keasey, K. (2014). A cross-country analysis of herd behavior in Europe. Journal of International Financial Markets, Institutions and Money, 32, 107-127. https://doi.org/10.1016/j.intfin.2014.05.008

Mossin, J. (1966). Equilibrium in a capital asset market. Econometrica: Journal of the econometric society, 43(4), 768-783. https://doi.org/10.2307/1910098

Ritter, J. R. (2003). Behavioral finance. Pacific-Basin finance journal, 11(4), 429-437. https://doi.org/10.1016/S0927-538X(03)00048-9

Roll, R. (1977). A critique of the asset pricing theory's tests. Part I: On past and potential testability of the theory. Journal of Financial Economics, 4(2), 129–176. https://doi.org/10.1016/0304-405X(77)90009-5

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. Journal of Economic Theory, 13(3), 341–360. https://doi.org/10.1016/0022-0531(76)90046-6

Roll, R., & Ross, S. A. (1980). An empirical investigation of the arbitrage pricing theory. The Journal of Finance, 35(5), 1073–1103. https://doi.org/10.1111/j.1540-6261.1980.tb02197.x

Scharfstein, D. S., & Stein, J. C. (1990). Herd behavior and investment. The American Economic Review, 465-479. Retrieved 25-9-2025 from https://www.jstor.org/stable/2006678

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. Journal of Finance, 19(3), 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? The American Economic Review, 71(3), 421–436. Retrieved 10-10-2025 from https://www.jstor.org/stable/1802789

Shiller, R. J. (2000). Irrational Exuberance. Princeton University Press.

Shleifer, A., & Summers, L. H. (1990). The noise trader approach to finance. Journal of Economic Perspectives, 4(2), 19-33. Retrieved 10-10-2025 from https://www.jstor.org/stable/1942888

Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. Journal of Finance, 52(1), 35–55. https://doi.org/10.2307/2329555

Sias, R. W. (2004). Institutional herding. The Review of Financial Studies, 17(1), 165-206. https://doi.org/10.1093/rfs/hhg035

Spyrou, S. (2013). Herding in financial markets: a review of the literature. Review of Behavioral Finance, 5(2), 175-194. https://doi.org/10.1108/RBF-02-2013-0009

Tan, L., Chiang, T. C., Mason, J. R., & Nelling, E. (2008). Herding behavior in Chinese stock markets: An examination of A and B shares. Pacific-Basin finance journal, 16(1-2), 61-77. https://doi.org/10.1016/j.pacfin.2007.04.004

Zhang, Y., & Zhao, R. (2004). The valuation differential between class A and B shares: country risk in the Chinese stock market. Journal of International Financial Management & Accounting, 15(1), 44-59. https://doi.org/10.1111/j.1467-646X.2004.00101.x

Appendices

Appendix 1. Market-wide herding analysis over the full sample period without an AR(1) specification

Robustness check: estimated coefficients for market-wide herding over the full sample period without the AR(1) specification

Market	γ_0	γ_1	γ_2	γ_3	Adj. R^2
India	0,0110***	0,0275**	0,2804***	0,3839	0,4010
	(41,1648)	(2,0908)	(6,7018)	(1,0658)	
China	0,0147***	-0,0034	0,1609***	-0,2581	0,1153
	(46,6950)	(-0,3243)	(5,8502)	(-0,4840)	

This table reports the estimated coefficients of equation (14) for the full sample period 02.07.2018-30.6.2022. A negative and statistically significant coefficient on γ_3 indicates herding. The T-statistics are reported in parentheses. HAC (Newey–West) standard errors are used to correct for heteroskedasticity and autocorrelation. Statistical significance is reported as follows:
* Statistical significance at the 10% level
**Statistical significance at the 5% level
***Statistical significance at the 1% level

Appendix 2. Market-wide herding analysis over the three subperiods without an AR(1) specification

Robustness check: estimated coefficients for market-wide herding over the three subperiods without the AR(1) specification

Panel A: Pre-COVID period

	γ_0	γ_1	γ_2	γ_3	Adj. R^2
Market					
India	0,0113***	-0,0033	0,1565***	4,8759***	0,3222
	(35,3310)	(-0,1548)	(2,9641)	(3,4748)	
China	0,0124***	-0,0201*	0,1567***	0,3843	0,2668
	(43,9904)	(-1,8462)	(4,6019)	(0,6588)	

Panel B: Outbreak period

	γ_0	γ_1	γ_2	γ_3	Adj. R^2
India	0,0146***	0,0493	0,3530***	-0,5000	0,5285
	(12,8408)	(1,6589)	(5,4502)	(-0,8706)	
China	0,0141***	0,0074	0,1781***	-0,4185	0,2280
	(24,4951)	(0,2124)	(2,7662)	(-0,3794)	

Panel C: Post-outbreak period

	γ_0	γ_1	γ_2	γ_3	Adj. R^2
India	0,0115***	-0,0074	0,1570***	0,3614	0,1702
	(39,6958)	(-0,4136)	(3,7515)	(0,2364)	
China	0,0167***	0,0127	0,1497***	0,2251	0,1035
	(40,2965)	(0,8360)	(3,5888)	(0,2147)	

This table reports the estimated coefficients of equation (14) for the three subperiods. A negative and statistically significant γ_3 result indicates herding. The T-statistics are reported in parentheses. HAC (Newey–West) standard errors are used to correct for heteroskedasticity and autocorrelation. The subperiods and statistical significance are as follows:

Pre-COVID period: 02.07.2018-29.1.2020

Outbreak period: 30.1.2020-30.6.2020

Post-outbreak period: 01.07.2020-30.6.2022

* Statistical significance at the 10% level

**Statistical significance at the 5% level

***Statistical significance at the 1% level

Appendix 3. Market-wide herding analysis during up- and down-market days over the full sample period without an AR(1) specification

Robustness check: estimated coefficients for market-wide herding during up- and down-market days over the full sample period without the AR(1) specification

	γ_0	γ_1	γ_2	γ_3	γ_4	Adj. R^2
Market						
India	0,0114***	0,1918***	-0,2501***	4,1007***	0,1903	0,4155
	(47,6589)	(3,8991)	(-6,7220)	(2,8723)	(0,5388)	
China	0,0147***	0,1683***	-0,1631***	-0,6595	-0,1911	0,1145
	(46,4407)	(5,0920)	(-5,204)	(-0,6330)	(-0,3203)	

This table reports the estimated coefficients of equation (15) for the full sample period 02.07.2018-30.6.2022. A negative and statistically significant coefficient on γ_3 indicates herding during up-market days while a negative and statistically significant coefficient on γ_4 indicates herding during down-market days . The T-statistics are reported in parentheses. HAC (Newey–West) standard errors are used to correct for heteroskedasticity and autocorrelation. Statistical significance is reported as follows:

* Statistical significance at the 10% level

**Statistical significance at the 5% level

***Statistical significance at the 1% level